

LOAN DOCUMENT

PHOTOGRAPH THIS SHEET

AD-A261 990



DTIC ACCESSION NUMBER

LEVEL

LEVEL

1

INVENTORY

AFDSR-TR-93-0112

DOCUMENT IDENTIFICATION

Dec 92

~~DISTRIBUTION STATEMENT~~

Approved for public release.
Distribution Unlimited

DISTRIBUTION STATEMENT

ACCESSION FOR	
NTIS	GRA&I
DTIC	TRAC
UNANNOUNCED	
JUSTIFICATION	
BY	
DISTRIBUTION/	
AVAILABILITY CODES	
DISTRIBUTION	AVAILABILITY AND/OR SPECIAL
A-1	

DISTRIBUTION STAMP

DTIC
ELECTE
MAR 10 1993
S C D

DATE ACCESSIONED

DATE RETURNED

DATE RETURNED

93 8 8 301
~~93 3 4 076~~

DATE RECEIVED IN DTIC

REGISTERED OR CERTIFIED NUMBER

93-04704



PHOTOGRAPH THIS SHEET AND RETURN TO DTIC-FDAC

H
A
N
D
L
E

W
I
T
H

C
A
R
E

UNITED STATES AIR FORCE

SUMMER RESEARCH PROGRAM -- 1992

**SUMMER FACULTY RESEARCH PROGRAM
(SFRP) REPORTS**

VOLUME 3

PHILLIPS LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES

**5800 UPLANDER WAY
CULVER CITY, CA 90230-6608**

SUBMITTED TO:

**LT. COL. CLAUDE CAVENDER
PROGRAM MANAGER**

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

**BOLLING AIR FORCE BASE
WASHINGTON, D.C.**

DECEMBER 1992

REPORT DOCUMENTATION PAGE

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
28 Dec 92

3. PERIODICITY
Annual 1 Sep 91 - 31 Aug 92

4. TITLE AND SUBTITLE

v-3

1992 Summer Faculty Research Program (SFRP)
Volumes 1 - 16

F49620-90-C-0076

5. AUTHOR
Mr Gary Moore

6. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Research & Development Laboratories (RDL)
5800 Uplander Way
Culver City CA 90230-6600

7. FUNDING MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFOSR/NI
110 Duncan Ave., Suite B115
Bldg 410
Bolling AFB DC 20332-0001
Lt Col Claude Cavender

8. SUPPLEMENTARY NOTES

9. DISTRIBUTION STATEMENT

UNLIMITED

10. ABSTRACT (Maximum 200 words)

The purpose of this program is to develop the basis for continuing research of interest to the Air Force at the institution of the faculty member; to stimulate continuing relations among faculty members and professional peers in the Air Force to enhance the research interests and capabilities of scientific and engineering educators; and to provide follow-on funding for research of particular promise that was started at an Air Force laboratory under the Summer Faculty Research Program.

During the summer of 1992 185 university faculty conducted research at Air Force laboratories for a period of 10 weeks. Each participant provided a report of their research, and these reports are consolidated into this annual report.

11. SUBJECT TERMS

12. SECURITY CLASSIFICATION OF REPORT

UNCLASSIFIED

13. SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

14. SECURITY CLASSIFICATION OF ABSTRACT

UNCLASSIFIED

UL

UNITED STATES AIR FORCE
SUMMER RESEARCH PROGRAM -- 1992
SUMMER FACULTY RESEARCH PROGRAM (SFRP) REPORTS

VOLUME 3

PHILLIPS LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES

5800 Uplander Way
Culver City, CA 90230-6608

Program Director, RDL
Gary Moore

Program Manager, AFOSR
Lt. Col. Claude Cavender

Program Manager, RDL
Billy Kelley

Program Administrator, RDL
Gwendolyn Smith

Submitted to:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Bolling Air Force Base

Washington, D.C.

December 1992

PREFACE

This volume is part of a 16-volume set that summarizes the research accomplishments of faculty, graduate student, and high school participants in the 1992 Air Force Office of Scientific Research (AFOSR) Summer Research Program. The current volume, Volume 3 of 16, presents the final research reports of faculty (SFRP) participants at Phillips Laboratory.

Reports presented herein are arranged alphabetically by author and are numbered consecutively -- e.g., 1-1, 1-2, 1-3; 2-1, 2-2, 2-3.

Research reports in the 16-volume set are organized as follows:

VOLUME	TITLE
1	Program Management Report
2	Summer Faculty Research Program Reports: Armstrong Laboratory
3	Summer Faculty Research Program Reports: Phillips Laboratory
4	Summer Faculty Research Program Reports: Rome Laboratory
5A	Summer Faculty Research Program Reports: Wright Laboratory (part one)
5B	Summer Faculty Research Program Reports: Wright Laboratory (part two)
6	Summer Faculty Research Program Reports: Arnold Engineering Development Center; Civil Engineering Laboratory; Frank J. Seiler Research Laboratory; Wilford Hall Medical Center
7	Graduate Student Research Program Reports: Armstrong Laboratory
8	Graduate Student Research Program Reports: Phillips Laboratory
9	Graduate Student Research Program Reports: Rome Laboratory
10	Graduate Student Research Program Reports: Wright Laboratory
11	Graduate Student Research Program Reports: Arnold Engineering Development Center; Civil Engineering Laboratory; Frank J. Seiler Research Laboratory; Wilford Hall Medical Center
12	High School Apprenticeship Program Reports: Armstrong Laboratory
13	High School Apprenticeship Program Reports: Phillips Laboratory
14	High School Apprenticeship Program Reports: Rome Laboratory
15	High School Apprenticeship Program Reports: Wright Laboratory
16	High School Apprenticeship Program Reports: Arnold Engineering Development Center; Civil Engineering Laboratory

1992 FACULTY RESEARCH REPORTS

Phillips Laboratory

<u>Report Number</u>	<u>Report Title</u>	<u>Author</u>
1	Coherent Heterodyne Array Doppler Imaging	Dr. Richard Anderson
2	A Brief Study of Passive Viscous Damping for the SPICE Bulkhead Structure	Dr. Joseph R. Baumgarten
3	Calibration Techniques for a Low Energy X-Ray Irradiation Chamber	Dr. Raymond D. Bellem
4	(Report not received)	
5	Ultrawideband Antennas with Low Dispersion	Dr. Albert W. Biggs
6	Optical Angle-Angle Doppler Imaging	Dr. J. K. Boger
7	Second-Harmonic Generation in Corona-Poled Materials	Dr. Gene O. Carlisle
8	Use of Optical Fibers in Long Baseline Interferometric Imaging	Dr. Douglas A. Christensen
9	Analysis of ONYNEX and MSRP Seismic Refraction Data in New England	Dr. John Ebel
10	Preparations for Neutron Scattering Investigations of Liquid-Crystal Polymers	Dr. David M. Elliott
11	Establishment of an Arcjet Optical Diagnostics Facility at Phillips Lab	Dr. Daniel A. Erwin
12	Ion-Molecule Reactions at High Temperatures	Dr. Jeffrey F. Friedman
13	Environmentally Safe Propellants: Ionophilic Polymer in Liquid Salt Systems	Dr. Daniel Lee Fuller
14	An Investigation of the Feasibility of Various Energetic Salt Combinations for Use in Solution Propellants	Dr. Vincent P. Giannamore
15	Development of a Prototype Lidar System at the Starfire Optical Range	Dr. Gary G. Gimmestad
16	Dynamics of the Reactions of Ar^+ , N_2^+ and Kr^+ with NO	Dr. Susan T. Graul
17	Radar Simulation and Analysis Tools for Satellite Analysis - A Recommendation	Dr. J. M. Henson
18	Optical and Atmospheric Turbulence	Dr. Mayer Humi
19	A Thermal Modeling of Directionally Woven Carbon-Carbon Fiber Materials - Prediction of Orthotropic Thermal Conductivity	Dr. Amir Karimi
20	Spin Dynamics of Lageos Satellite	Dr. Arkady Kheyfets

Phillips Laboratory (cont'd)

<u>Report Number</u>	<u>Report Title</u>	<u>Author</u>
21	Generation of ELF and VLF Waves by a Thermal Instability Excited in the HF Heater-Modulated Polar Electrojet	Dr. S. P. Kuo
22	(Report not received)	
23	Equation of Radiative Energy Transfer in Electromagnetic Theory with Wave-Optical Contributions	Dr. Arvind S. Marathay
24	Ion-Molecule Reactions at High Temperatures	Dr. Thomas M. Miller
25	Particle Simulations of Plasmas	Dr. R. D. Murphy
26	Pointing and Accuracy Analysis of Solar Concentrators	Dr. Rupa Purasinghe
27	Mixed-Mode Fracture of Solid Propellants	Dr. Krishnaswamy Ravi- Chandar
28	Vibration and Compression Testing of Composite Isogrid Panels	Dr. Christopher A. Rotz
29	Solar Scintillation and the Monitoring of Solar Seeing	Dr. Edward J. Seykora
30	Space Debris Impact Effects on Spacecraft Materials	Dr. Jon H. Shively
30A	Estimation of Misalignment Parameters of Multi-Aperture Telescopes	Dr. Johanna K. H. Stenzel
31	Satellite Fragmentation Due to Explosion and Collision	Dr. A. Tan
32	Spectroscopic Data of Atmospheric Interest	Dr. R. H. Tipping
33	The Role of Atomization in Liquid Propellant Rocket Combustion Instability	Dr. Robert J. Turnbull
34	Experimental Investigation of Homogeneous and Heterogeneous Nucleation/Condensation Processes and Products in Coil	Dr. Philip D. Whitefield
35	Installation of Oscillating Neural Network Algorithms into a Khoros Toolbox	Dr. Carla C. S. Williams
36	The Trajectories of Eastern Hemisphere Solar Electron Events as Inferred from ISEE-3 Radio and Particle Data	Dr. Robert F. Willson

COHERENT HETERODYNE ARRAY DOPPLER IMAGING

Richard Anderson
Professor
Department of Physics
University of Missouri-Rolla
Rolla, MO 65401

Final Report for:
Summer Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

September 1992

COHERENT HETERODYNE ARRAY DOPPLER IMAGING

Richard Anderson
Professor
Department of Physics
University of Missouri

Abstract

In initial research performed previously coherent heterodyne angle-angle (2-D) imaging was performed. In this research angle-angle-Doppler imaging will be performed. In future research ranging shall be added so angle-angle-range and 3-D-Doppler data may be acquired. The theory of the phase-up and of angle-angle-Doppler imaging is presented. An experiment will be performed on angle-angle-Doppler imaging.

INTRODUCTION

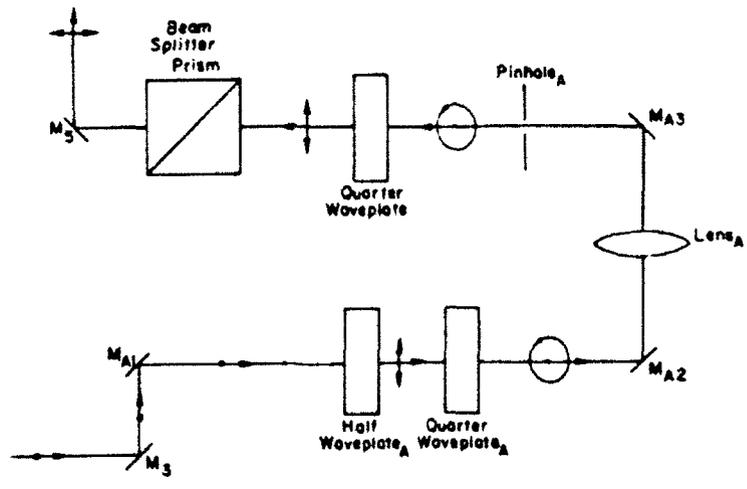
There has been only one previous experiment on coherent heterodyne angle-angle-Doppler array imaging¹⁻². No attempt was possible in these studies to average speckle realizations and to attempt to recover an approximate incoherent angle-angle-Doppler image of the target. The study described above was performed with small and very slow motion of the target and because of the slowness of the camera no signal averaging was possible. The image was improved by controlling the incident beam intensity.

There have been several studies on coherent angle-angle and angle-angle-Doppler imaging employing a single detector and a 2-D scanned beam³⁻⁵. The speckled image is the composite of all points in the scan. Measurements of the line-of-sight or radial speed gives rise to the Doppler frequency shift in the target signal and by measuring the Doppler shift at various target positions an angle-angle-Doppler image was presented.

THEORY AND DISCUSSION

a. Theory of Phase-up of the System

A basic reflective coherent heterodyne array imaging system is shown in Figure 1. A heterodyne system is a form of the Mach-Zehnder interferometer generally with different arm lengths. The source is a laser and two beams are produced by a beam splitter with the signal beam in a laboratory system generally 1/10 th the intensity of the local oscillator (LO) beam. Both beams pass through an acousto-optic modulator (AOM) which shifts the frequency so the mixed beams have a difference frequency which allows each pixel in the pupil plane of the camera to be sampled adequately to resolve the heterodyne signal in each frame (speckle realization). The frequency difference between the two beams should be at least 1/10 the camera scan speed so each heterodyne waveform is sampled 10 times in each frame. In the phase-up procedure the next element in LO beam is a spatial filter followed by a pinhole. The signal beam is focussed on a pinhole mounted on a mirror



Alternate method to produce target pinhole source

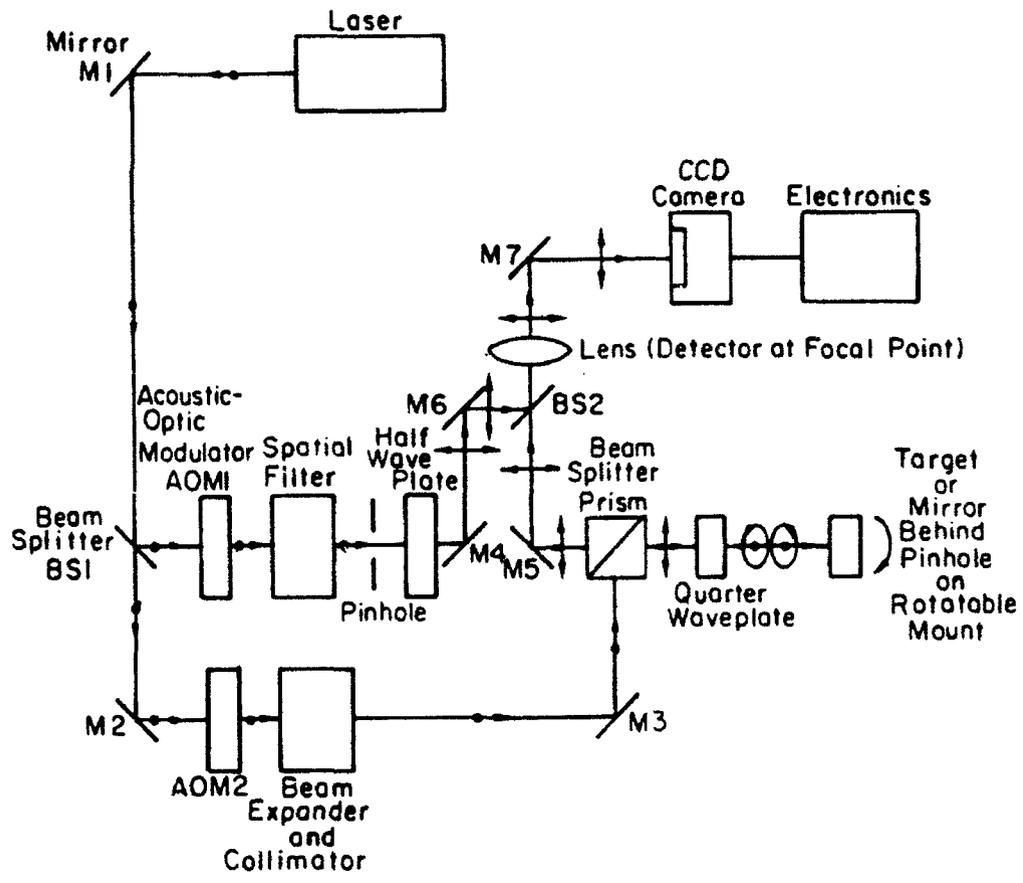


Figure 1 Sketch of Apparatus with Reflective Target

or a small bearing mounted on a black surface. (An alternate method using a pinhole in the place of the target is shown in the inset. The phase-up is from the pinhole to the detector through any distorting medium). This focussed beam is reflected from a prism beam splitter and passes through a quarter waveplate producing circularly polarized light. The specularly reflected light from the pinhole or bearing is polarized in the opposite sense and on passing through the quarter waveplate it is polarized in the plane of the paper and is transmitted by the prism beam splitter. The polarization of the LO beam is rotated by 90° by passing through a half waveplate.

When the target replaces the pinhole in the signal arm, the spatial filter is replaced by a beam expander and collimator which flood illuminates the entire target and the target is in the same position as the previous pinhole. The real target is a diffuse, random reflective target with a definite shape. In the phase-up procedure both the LO and signal targets are pinholes. A phase error between the two beams arises from the pinholes to the detector. It arises from the lack of centering of optical components, poor quality mirrors, the fact the optical surfaces of the beam splitter are not perfect and one beam is reflected from and the other transmitted through the splitter, aberrations in the lens, and phase differences between detector elements. The introduction of a distorting medium in one path could be used to simulate atmospheric turbulence. Clark and Desandre¹ have discussed this problem for a homodyne system.

In the phase-up procedure the target is a pinhole as is the LO target. The emergent wavefront are diverging spherical waves that are intercepted by the lens and the Fourier transform is produced in the focal plane of the lens which is the pupil plane of the camera. The transform of the field of a small pinhole is a Sombrero function which has a very broad central maximum and covers an area much larger than the area of the small CCD detector and its value is taken as unity over the detector area and U_p and U_{LO} are field amplitudes. The fields at a position vector \vec{r} of coordinates (X,Y) in the pupil plane of the detector are

$$U_p(\bar{r}, t) = U_p \text{Sombero}(\bar{r}/\lambda f) e^{j\phi'(\bar{r})} e^{j\omega_s t} \approx U_p e^{j[\omega_s t - \phi'(\bar{r})]} \quad (1)$$

and

$$U_{LO}(\bar{r}, t) = U_{LO} \text{Sombero}(\bar{r}/\lambda f) e^{j\phi''(\bar{r})} e^{j\omega_{LO} t} \approx U_{LO} e^{j[\omega_{LO} t - \phi''(\bar{r})]} \quad (2)$$

where $\phi'(\bar{r})$ and $\phi''(\bar{r})$ are the phase errors introduced by the respective paths from the pinholes through the detector. The angular frequencies of the signal and LO beams are ω_s and ω_{LO} , respectively. The subscripts LO and p indicate the local oscillator beam and the signal beam when the target is the pinhole. The total mixed field at the position vector \bar{r} in the pupil plane of the detector is

$$U_{\text{Total}}(\bar{r}, t) = U_p(\bar{r}, t) + U_{LO}(\bar{r}, t) \quad (3)$$

and the signal is

$$\begin{aligned} S_{p,LO}(\bar{r}, t) &= \eta \left\{ |U_p(\bar{r}, t)|^2 + |U_{LO}(\bar{r}, t)|^2 \right. \\ &\quad \left. + U_{LO} U_p \left[e^{j[\omega_{IF} t + \phi(\bar{r})]} + e^{-j[\omega_{IF} t + \phi(\bar{r})]} \right] \right\} \\ &= \eta \left\{ U_p^2 + U_{LO}^2 + U_{LO} U_p \left[e^{j[\omega_{IF} t + \phi(\bar{r})]} + e^{-j[\omega_{IF} t + \phi(\bar{r})]} \right] \right\} \\ &= \eta \left\{ U_p^2 + U_{LO}^2 + U_{LO} U_p \text{Cos}[\omega_{IF} t + \phi(\bar{r})] \right\} \quad (4) \end{aligned}$$

where $\phi(\bar{r}) = \phi'(\bar{r}) - \phi''(\bar{r})$ and $\omega_{IF} = \omega_s - \omega_{LO}$. The signal intensity is small compared to the LO intensity and can be neglected in comparison and the LO signal determines the shot noise limit of detection. The heterodyne signal is

$$S_{p,LO,het}(\bar{r}, t) = 2\eta U_{LO} U_p \text{Cos}[\omega_{IF} t + \phi(\bar{r})] \quad (5)$$

Equation (5) shall be written in a different form below:

$$S_{p,LO,het}(\bar{r}, t) = 2\eta U_p U_{LO} [\text{Cos}\omega_{IF} t \text{Cos}\phi(\bar{r}) - \text{Sin}\omega_{IF} t \text{Sin}\phi(\bar{r})] \quad (6)$$

The temporal Fourier transform yields signal components at ω_{IF} and $-\omega_{IF}$ and the dc signal at $\omega = 0$ so the heterodyne components are

$$S_{p,LO,het}(\bar{r}, \omega_{IF}) = \eta' U_p U_{LO} [\cos\phi(\bar{r}) - j \sin\phi(\bar{r})] = \eta' U_p U_{LO} e^{j\phi(\bar{r})} \quad (7a)$$

and

$$S_{p,LO,het}(\bar{r}, -\omega_{IF}) = \eta' U_p U_{LO} [\cos\phi(\bar{r}) - j \sin\phi(\bar{r})] = \eta' U_p U_{LO} e^{-j\phi(\bar{r})} \quad (7b)$$

Then

$$2\eta' U_p U_{LO} \cos\phi(\bar{r}) = S_{p,LO,het}(\bar{r}, \omega_{IF}) + S_{p,LO,het}(\bar{r}, -\omega_{IF}) \quad (8)$$

and

$$2\eta' U_p U_{LO} \sin\phi(\bar{r}) = j [S_{p,LO,het}(\bar{r}, -\omega_{IF}) - S_{p,LO,het}(\bar{r}, \omega_{IF})]$$

so

$$\tan\phi(\bar{r}) = \frac{j [S_{p,LO,het}(\bar{r}, -\omega_{IF}) - S_{p,LO,het}(\bar{r}, \omega_{IF})]}{[S_{p,LO,het}(\bar{r}, \omega_{IF}) + S_{p,LO,het}(\bar{r}, -\omega_{IF})]} \quad (9)$$

Thus the phase error can be determined pixel by pixel in the pupil plane of the camera.

b. Phasing-up a Target at Rest

The pinhole target is replaced by a real target which might be a small, diffuse, random reflective target of definite contour. The scattering points have a slightly different position if the target is rotated and the scattering gives a new speckle realization. The target is rotated by a stepping motor and each position is a different speckle realization of the target. Each position of the target is called a frame and data are taken while the target is at rest between rotations. The total motion of the target is small so the perspective of the target does not change appreciably. The intermediate frequency of the heterodyne signal is at least 1/10 the scan speed of the camera so in each frame the heterodyne wave form can reasonable be recovered at each pixel in the camera pupil plane. The transformed signals in these frames are averaged to reduce the effects of speckle. Since the target is at rest in each frame, there is no Doppler shift of the target signal.

In the LO arm the target is the same pinhole which again gives rise to a broad Sombero function whose amplitude will be considered

unity over the CCD camera pupil plane. The target field in the pupil plane of the detector is the diffraction pattern of the reflective target. The signals in the pupil plane at position vector \bar{r} are

$$U_T(\bar{r}, t) = kU_T(\bar{r}/\lambda f) e^{j\phi'(\bar{r})} e^{j\omega_s t}$$

and

$$(10)$$

$$U_{LO}(\bar{r}, t) = U_{LO} e^{j\phi''(\bar{r})} e^{j\omega_{LO} t}$$

so the mixed field in the pupil plane of the camera at position vector \bar{r} at time t is

$$U_{Total}(\bar{r}, t) = U_T(\bar{r}, t) + U_{LO}(\bar{r}, t) \quad (11)$$

The signal at this point is

$$S_{T,LO}(\bar{r}, t) = \left\{ \eta' |U_T(\bar{r}, t)|^2 + \eta |U_{LO}(\bar{r}, t)|^2 \right. \\ \left. - \eta'' U_{LO} U_T(\bar{r}/\lambda f) e^{j[\omega_{IF} t + \phi(\bar{r})]} + \eta'' U_{LO} U_T^*(\bar{r}/\lambda f) e^{-j[\omega_{IF} t - \phi(\bar{r})]} \right\}$$

$$(12)$$

where the LO field amplitude is constant and real. The first term is the dc target signal which is small and can be neglected compared to the LO signal. The second term is the dc LO signal and determines the shot noise limit of the system. The heterodyne signal is

$$S_{T,LO,heter}(\bar{r}, t) = 2\eta'' U_{LO} \left\{ U_{T,RI}(\bar{r}/\lambda f) \cos[\omega_{IF} t + \phi(\bar{r})] \right. \\ \left. - U_{T,Im}(\bar{r}/\lambda f) \sin[\omega_{IF} t + \phi(\bar{r})] \right\}$$

$$(13)$$

In order to retain the phase information about the target equation (12) should be written in the form

$$S_{T,LO,heter}(\bar{r}, t) = \eta'' U_{LO} \left\{ U_T(\bar{r}/\lambda f) e^{j\phi(\bar{r})} \left[\cos\omega_{IF} t + j\sin\omega_{IF} t \right] \right. \\ \left. - U_T^*(\bar{r}/\lambda f) e^{-j\phi(\bar{r})} \left[\cos\omega_{IF} t - j\sin\omega_{IF} t \right] \right\}$$

$$(14)$$

The temporal Fourier transform of this equation heterodyne signal gives single components at ω_{IF} and $-\omega_{IF}$ or

$$S_{T,LO,het}(\bar{r}, \omega_{IF}) = 1/2 \chi' U_{LO} \left[U_T(\bar{r}/\lambda f) e^{j\phi(\bar{r})} - U_T^*(\bar{r}/\lambda f) e^{-j\phi(\bar{r})} - U_T(\bar{r}) e^{j\phi(\bar{r})} - U_T^*(\bar{r}) e^{-j\phi(\bar{r})} \right] = \chi' U_{LO} U_T(\bar{r}/\lambda f) e^{j\phi(\bar{r})}$$

(15a)

and

$$S_{T,LO,het}(\bar{r}, -\omega_{IF}) = 1/2 \chi' U_{LO} \left[U_T(\bar{r}/\lambda f) e^{j\phi(\bar{r})} - U_T(\bar{r}/\lambda f) e^{j\phi(\bar{r})} - U_T^*(\bar{r}/\lambda f) e^{-j\phi(\bar{r})} - U_T^*(\bar{r}/\lambda f) e^{-j\phi(\bar{r})} \right]$$

$$= \chi' U_{LO} U_T^*(\bar{r}/\lambda f) e^{-j\phi(\bar{r})}$$

(15b)

The signal at angular frequency ω_{IF} in equation (15a) is multiplied by the signal in equation (7b) and at angular frequency $-\omega_{IF}$ the signal in equation (15b) is multiplied by the signal in equation (7a) and this yields the two equation

$$S_T(\bar{r}, \omega_{IF}) = S_{T,LO,het}(\bar{r}, \omega_{IF}) S_{p,LO,het}(\bar{r}, -\omega_{IF}) = \chi'' U_{LO}^2 U_p U_T(\bar{r}/\lambda f)$$

and

(16)

$$S_T(\bar{r}, -\omega_{IF}) = S_{T,LO,het}(\bar{r}, -\omega_{IF}) S_{p,LO,het}(\bar{r}, \omega_{IF}) = \chi'' U_{LO}^2 U_p U_T^*(\bar{r}/\lambda f)$$

The phase error is eliminated and the two resulting signals are proportional to the signal and its complex conjugate at position \bar{r} in the pupil plane of the detector for χ'' . U_{LO} and U_p are assumed constants.

The spatial Fourier transform is taken of these signals to determine the complex target field at a position vector \bar{u} of coordinates (x,y) in the target plane. The transformed signals are

$$S_T(\bar{u}, \omega_{IF}) = \chi'' U_{LO}^2 U_p U_T(\bar{u})$$

and

(17)

$$S_T(\bar{u}, -\omega_{IF}) = \chi'' U_{LO}^2 U_p U_T^*(\bar{u})$$

Assume that the target and the pupil plane of the camera lie at the conjugate focal planes of the lens, then the Fourier transform of the pupil plane image into the target plane image is

$$U_T(\bar{u}) = e^{2jkf} / j\lambda f \int_{\substack{\text{camera} \\ \text{pupil plane}}} U_T(\bar{r}'/\lambda f) e^{j(k\bar{r}' \cdot \bar{u}/f)} d^2\bar{r}'$$

where $U_T(\bar{r}'/\lambda f)$ is the signal field at point \bar{r}' in the pupil plane of the camera and $U_T(\bar{u})$ is target field at point \bar{u} in the target plane. The integral is a 2-D Fourier transform over only the specific propagation vector that determines the point \bar{u} in the target plane or over the propagation vector $\bar{k} = kx/\hat{f}\hat{i} + ky/\hat{f}\hat{j}$. The lens collects all allowed propagation vectors that had passed through it and it effectively integrates the signal over all allowed propagation vectors and the integral above is in reality a complete Fourier transform of the field from the pupil plane of the camera to the target plane.

The real and imaginary components of the field in the target plane are

$$m_{RI}(\bar{u}) = \left[S(\bar{u}, \omega_{IF}) + S(\bar{u}, -\omega_{IF}) \right] = 2x'' U_{LO}^2 U_p U_T(\bar{u}) \cos\theta(\bar{u})$$

and

$$m_{Im}(\bar{u}) = j \left[S(\bar{u}, -\omega_{IF}) - S(\bar{u}, \omega_{IF}) \right] = 2\eta''^2 U_{LO}^2 U_p U_T(\bar{u}) \sin\theta(\bar{u})$$

(18)

The phase angle between the components is

$$\tan\theta(\bar{u}) = m_{Im}(\bar{u})/m_{RI}(\bar{u})$$

(19)

The speckled intensity in the target plane is proportional to

$$S_T(\bar{u}, \omega_{IF}) S_T(\bar{u}, -\omega_{IF}) = x''^2 U_{LO}^4 U_p^2 |U_T(\bar{u})|^2 = x''^2 U_{LO}^4 U_p^2 I_{Ts}(\bar{u})$$

(20)

If the motion of the target between frames is small, a different speckle realization is measured. If the total motion over all frames is still reasonably small the same target perspective is measured in all frames and the intensities at each pixel may be averaged to eliminate the effects of speckle and obtain a representation of the incoherent image. In reality it requires the average to be made over an infinite number of frames to recover the true incoherent image. Thus,

$$I_{\text{incoherent}}(\bar{u}) \approx \sum_{i=1}^N S_T(\bar{u}, \omega_{IF}) S_T(\bar{u}, -\omega_{IF}) / N = K \sum_{i=1}^N I_{Ts}(\bar{u}) / N$$

(21)

The number of frames is N . If $N \Rightarrow \infty$, the true incoherent intensity is measured

$$I(\bar{u})_{\text{incoherent}} = \lim_{N \Rightarrow \infty} \frac{\sum_{i=1}^N S_T(\bar{u}, \omega_{IF}) S_T(\bar{u}, -\omega_{IF})}{N} = \lim_{N \Rightarrow \infty} \frac{\sum_{i=1}^N I_{TS}(\bar{u})}{N} \quad (22)$$

c. Theory to Retrieve Angle-Angle-Doppler Images

Heterodyne detection⁶ is used for three basic reasons: (1) It is used to retrieve very weak repetitive signals from noise. (2) It is used to obtain spatial phase of the target. (3) It is used to obtain the real and imaginary components of the field in target space and as a result the speckled target image. If speckle averaging is possible, the incoherent target image can be approximated. In a cw repetitive experiment it is the preferred technique in order to recover a weak signal from noise. It retains the most information about the target for it gives the complex target field and its complex conjugate, the phase between the real and imaginary components of the target field and the real and imaginary field components. It measures the target field instead of intensity and the heterodyne signals are directly proportional the complex target field and its complex conjugate in detector space. Heterodyne detection approaches the quantum limit of detection (shot noise) when in the detector plane, $U_T \Rightarrow 0$. In heterodyne detection a strong LO field is mixed with the weak signal field from the target and the mixed field is detected in the detector plane. It is generally assumed that the LO field has been shaped so it has a real, constant amplitude in the detector plane. The detector measures an intensity with four components: the target intensity, the LO intensity, and two interference intensity terms at the intermediate frequency between the signal and LO fields. The first interference term has an amplitude directly proportional to the complex target field in the detector plane and the other term is directly proportional to the complex conjugate of this field. If it is possible, a heterodyne system is operated near the quantum noise limit or $I_T \ll I_{LO}$. (This condition is easily fulfilled for an object in space which is

illuminated from a ground based laser and the resultant signal intensity is recorded in the detector plane.) Then the LO intensity determines the noise limit of detection and the product heterodyne terms have a reasonable amplitudes so they may be detected. The heterodyne terms do not exhibit the true speckle of the target. This is observed in the target plane for a single speckle realization (frame) when the complex target field is taken times its complex conjugate. These speckle effects are averaged out by averaging the target intensity in the target plane over many different speckle realizations.

Some assumptions must be made about the target. It will be assumed to be a diffuse, random scatterer where the coherent, deterministic character which gives rise to "glints" (specular reflections) and retroreflection are neglected. This diffuse reflectance is referred to as speckle reflectivity for it gives rise to incoherent scattering and interference effects in the final image. It will be assumed that this speckle reflection coefficient $\rho_s(\bar{u})$ is a circulo-complex Gaussian function whose time ensemble averaged moments obey the properties below:

$$\begin{aligned} \langle \rho_s(\bar{u}_i) \rangle &= 0 \\ \langle \rho_s(\bar{u}_i) \rho_s(\bar{u}_j) \rangle &= 0 \end{aligned} \quad (23)$$

and

$$\langle \rho_s(\bar{u}_i) \rho_s^*(\bar{u}_j) \rangle = \lambda^2 \mathbb{R}_s(\bar{u}_i) \delta(\bar{u}_i - \bar{u}_j) / \pi$$

where $\mathbb{R}_s(\bar{u}_i)$ is the target reflectance at target position \bar{u}_i in the target plane and equals $|\rho_s(\bar{u}_i)|^2$. These equations are statements that a spatially coherent incident laser beam is scattered as a spatially incoherent beam. The δ function implies the scattering is non-directional. The scattered signal still has a temporal coherence and constructive and destructive interference effects occur in the detector giving a speckled image⁷.

In heterodyne detection after the temporal and spatial Fourier transforms are taken, it is the complex field in the target plane that is measured. This field is the product of the target reflection coefficient $\rho_s(\bar{u})$ times the complex incident laser field on the target $U_i(\bar{u})$ so the scattered field is $U_T(\bar{u}) = \rho_s(\bar{u})U_i(\bar{u})$. For a space based

target the incident field from the telescope in the up-link through the atmosphere would be distorted and in the target plane at each point \bar{u} its phase would be different and also its amplitude may be changed so at the target $U_i(\bar{u}) = U_i'(\bar{u})e^{j\theta'(\bar{u})}$ where U_i' is the amplitude of U_i and θ' is the phase change. The reflection coefficient is a complex number and since our target is diffuse and no deterministic surface profile changes are assumed, it is $\rho_s(\bar{u}) = \rho_s'(\bar{u})e^{j\theta''(\bar{u})}$ where ρ_s' is its amplitude and θ'' is the phase variation produced by the surface microstructure. The scattered field $U_T(\bar{u}) = \rho_s'(\bar{u})U_i'(\bar{u})e^{j[\theta'(\bar{u})+\theta''(\bar{u})]} = \rho_s'(\bar{u})U_i'(\bar{u})e^{j\theta(\bar{u})} = U_T'(\bar{u})e^{j\theta(\bar{u})}$ where θ is the resultant phase change and U_T' is the amplitude of the scattered field. In angle-angle and angle-angle-Doppler imaging the intensity is measured and it is the target field time its complex conjugate so the uplink and target phase cancel.

The target is a random, diffuse, reflective rotating object. All points an equal distance from the axis of rotation have the same line-of-sight speed, v_{\parallel} . The target motion must be small and one microscopic surface feature must not be replaced by another. If this happens, one is sampling signals over more than one speckle realization and the time average is zero from equation (23). This is equivalent to stating that the motion must be a small and slow during one speckle realization so the transverse motion of a speckle is a small fraction of its size so there will not be a smearing of the image in the detector plane.

The speckle size is approximately $w_s = \lambda R/W$ where λ is the wavelength, R is the range, and W is the target size. For a rotating target with $W \cong 1$ cm, $R \cong 1$ m, and $\lambda = 0.5$ or $1 \mu\text{m}$ then $w_s \cong 50$ or $100 \mu\text{m}$, respectively. The Doppler shift for a reflective target is $f_D = 2fv_{\parallel}/c = 2v_{\parallel}/\lambda = k_D v_{\parallel}$ where $k_D = 4$ or 2 MHz/m/s at $\lambda = 0.5$ or $1 \mu\text{m}$. This means a reflective target must move at 0.25 or $0.5 \mu\text{m/s}$, respectively, for a Doppler frequency shift of 1 Hz. For the rotating target the maximum transverse speed $v_{\perp\text{max}}$ equals $v_{\parallel\text{max}}$ so if a frame is sampled at 0.1 s the transverse motion is 0.025 or $0.05 \mu\text{m}$ or the target is

effectively at rest.

For a rotating target of 0.5 cm radius the rotational speeds are 0.5×10^{-5} or 1×10^{-5} radian/s at $\lambda = 0.5$ or $1 \mu\text{m}$. An approximate calculation gives a rotational speed of $\approx 2 \times 10^{-5}$ radian/s for a mount in the laboratory and this mount might be marginally useful. The important factor in all of these calculations is the quantity $k_D = 2/\lambda$ for it dictates in our experiment that the motion must be slow or the sampling speed must be fast. Our camera sampling speed restricts our experiment to slow line-of-sight speeds.

In real target acquisition in space the transverse motion of the speckle is eliminated by very accurate tracking of the target. Fine high resolution tracking is the most important requirement in active imaging⁸. If adequate tracking is achieved, for a target at ≈ 100 km the angular sweep speed is $\approx 1^\circ/\text{s}$ and the target can be followed without transverse displacement of the speckle.

Let us designate the point in the rotating target associated with a Doppler shift ω_{Di} by the subscript i on the target position coordinate and $\bar{u} = \bar{u}_i$. These points with a constant v_{\parallel} are always at the same position in the target, but the target is continually rotating so in order to measure a single speckle realization the measurement must be made in a time interval so the target motion is negligible. This can be done by making the measurement in a very short period of time for a moving target in the m/s range or as in our case the target rotational speed can be made very slow and the measurement is made in tenths of seconds. For convenience of writing the theoretical equations the target and the LO pinhole are placed at the conjugate focal planes of the transform lens from the detector. Then at the position vector \bar{r} in the pupil plane of the detector with coordinates (X,Y) the target field is

$$\begin{aligned}
 U_T(\bar{r}, t) &= e^{2jkf/j\lambda f} \int U_T(\bar{u}_i) e^{-jk\bar{u}_i \cdot \bar{r}/f} d^2\bar{u}_i e^{j(\omega_s \pm \omega_D)t} \\
 &= \left[e^{jkf/j\lambda f} \int_{\text{target}} U_T(\bar{u}_i) e^{-jk\bar{u}_i \cdot \bar{r}/f} d^2\bar{u}_i \right] e^{jkf} e^{j(\omega_s \pm \omega_D)t}
 \end{aligned}$$

$$= U_T(\bar{r}/\lambda f) e^{jkf} e^{j(\omega_s \pm \omega_D)t} = K U_T(\bar{r}/\lambda f) e^{j(\omega_s \pm \omega_D)t} \quad (24)$$

The integral focusses to the point \bar{r} propagation vectors from all target points having the direction given by the vector $\bar{k} = k\bar{r}/f$ and each target point propagates such a vector if it is a diffuse target. All possible Doppler frequency shifts are focussed at the point \bar{r} . The lens collects all allowed propagation vectors determined by its size and forms a complete 2-D Fourier transformed image of the target in the pupil plane of the camera. Equation (24) above must be corrected for any phase shifts produced by the optical system and any introduced distorting medium from the target through the pupil plane of the camera and this phase error has been previously designated $\phi'(\bar{r})$ so

$$U_T(\bar{r}, t) = K U_T(\bar{r}/\lambda f) e^{j[(\omega_s \pm \omega_D)t + \phi'(\bar{r})]} \quad (25)$$

Again this signal is mixed with the LO signal where the LO beam target is a pinhole and is given in equation (10). The mixed field at the detector is

$$U_{\text{Total}}(\bar{r}, t) = U_T(\bar{r}, t) + U_{\text{LO}}(\bar{r}, t) \quad (26)$$

The signal at position vector \bar{r} in the pupil plane of the detector is

$$\begin{aligned} S_{\text{T,LO}}(\bar{r}, t) &= \eta'' |U_T(\bar{r}, t)|^2 + \eta |U_{\text{LO}}(\bar{r}, t)|^2 \\ &+ \eta' U_{\text{LO}} U_T(\bar{r}/\lambda f) e^{j[(\omega_{\text{IF}} \pm \omega_D)t + \phi(\bar{r})]} \\ &+ \eta' U_{\text{LO}}^* U_T^*(\bar{r}/\lambda f) e^{j[(\omega_{\text{IF}} \pm \omega_D)t + \phi(\bar{r})]} \end{aligned} \quad (27)$$

where $\omega_{\text{IF}} = \omega_s - \omega_{\text{LO}}$ and $\phi(\bar{r}) = \phi'(\bar{r}) - \phi''(\bar{r})$. If the heterodyne condition is met, then $\eta'' |U_T(\bar{r}, t)|^2 \ll \eta |U_{\text{LO}}(\bar{r}, t)|^2$ and $\eta' U_{\text{LO}} |U_T(\bar{r}/\lambda f)|$ so the LO intensity dominates and determines the noise limit of detection. The heterodyne signal at Doppler shift ω_D is

$$\begin{aligned} S_{\text{T,LO,heter}}(\bar{r}, t) &= \eta' U_{\text{LO}} U_T(\bar{r}/\lambda f) e^{j[(\omega_{\text{IF}} \pm \omega_D)t + \phi(\bar{r})]} \\ &- \eta' U_{\text{LO}}^* U_T^*(\bar{r}/\lambda f) e^{-j[(\omega_{\text{IF}} \pm \omega_D)t + \phi(\bar{r})]} \end{aligned} \quad (28)$$

If the temporal Fourier transform is taken of equation (30), the

two Doppler broadened signals about ω_{IF} and $-\omega_{IF}$ are measured. Their width ranges from $\omega_{IF} - \omega_{Dmax}$ to $\omega_{IF} + \omega_{Dmax}$ and $-(\omega_{IF} - \omega_{Dmax})$ to $-(\omega_{IF} + \omega_{Dmax})$ where ω_{Dmax} is the maximum Doppler shift in either direction.

If the temporal transform is taken of the signal in the detector plane, the Doppler broadened signal is measured. This signal can be divided into small frequency bins of width $\Delta\omega_D \ll \omega_{Dmax}$. The spatial transform of these signals in each bin will locate a spot in the target plane where the line-of-sight speed $v_{||}$ is nearly constant with a spread of $\Delta v_{||}/2$. This technique might have aliasing problems near $\pm \omega_{Dmax}$ where the signal is small and it might show interference effects near $\omega_D = 0$.

It should be noted that $U_T(\bar{r}/\lambda f) = U_T'(\bar{r})e^{j\gamma(\bar{r})}$
 $= \kappa U_T(\bar{r}/\lambda f) * F\left\{e^{j\theta(\bar{u})}\right\}_{\bar{u}=\bar{r}/\lambda f}$ where $*$ is the convolution and F is the spatial transform of the bracketed function into the detector plane and
 $U_T^*(\bar{r}/\lambda f) = U_T'(\bar{r}/\lambda f)e^{-j\gamma(\bar{r})} = \kappa U_T'(\bar{r}/\lambda f) * F\left\{e^{-j\theta(\bar{u})}\right\}_{\bar{u}=\bar{r}/\lambda f}$. When the temporal Fourier transform is taken, then signal in the bin where the Doppler shift is ω_D is

$$\begin{aligned} S_T(\bar{r}, \omega_{IF} + \omega_D) &= \kappa' U_{LO} U_T(\bar{r}/\lambda f) e^{j\phi(\bar{r})} \\ &= \kappa' U_{LO} U_T'(\bar{r}/\lambda f) * F\left\{e^{j\theta(\bar{u})}\right\}_{\bar{u}=\bar{r}/\lambda f} e^{j\phi(\bar{r})} \end{aligned} \quad (29a)$$

and

$$\begin{aligned} S_T[\bar{r}, -(\omega_{IF} + \omega_D)] &= \kappa' U_{LO} U_T^*(\bar{r}/\lambda f) e^{-j\phi(\bar{r})} \\ &= \kappa' U_{LO} U_T'(\bar{r}/\lambda f) * F\left\{e^{-j\theta(\bar{u})}\right\}_{\bar{u}=\bar{r}/\lambda f} e^{-j\phi(\bar{r})} \end{aligned} \quad (29b)$$

Equation (29a) is multiplied by the phase-up equation (7b) and equation (29b) is multiplied by equation (7a) to remove the phase error in the signals. This gives

$$S_T(\bar{r}, \omega_{IF} + \omega_D) = \kappa'' U_{LO}^2 U_p U_T(\bar{r}/\lambda f) \quad (30a)$$

and

$$S_T[\bar{r}, -(\omega_{IF} + \omega_D)] = \kappa'' U_{LO}^2 U_p U_T^*(\bar{r}/\lambda f) \quad (30b)$$

Assuming the target and pinhole and the detector are at conjugate focal planes of the lens, then the spatial Fourier transform is of the form

$$\begin{aligned}
 U_T(\bar{u}) &= e^{2jkf} / j\lambda f \int U_T(\bar{r}' / \lambda f) e^{-j\{k\bar{r}' \cdot \bar{u} / f\}} d^2\bar{r}' = U_T(\bar{u}) e^{jkf} \\
 &\quad \text{camera} \\
 &\quad \text{pupil plane} \\
 &= kU_T(\bar{u}) e^{j\theta(\bar{u})} = k \rho'_S(\bar{u}) U'_I(\bar{u}) e^{j\theta(\bar{u})} \quad (31)
 \end{aligned}$$

The spatial transform of equation (30a) and (30b) give signals at the points in the target plane moving at speed $v_{\parallel} \pm \Delta v_{\parallel} / 2$ where $\Delta v_{\parallel} = \Delta\omega_D / 2\pi k_D$ so

$$\begin{aligned}
 m_{-, \omega_D}(\bar{u}_i) &= K U_T(\bar{u}_i) \\
 \text{and} & \quad (32) \\
 m_{-, \omega_D}(\bar{u}_i) &= K U_T^*(\bar{u}_i)
 \end{aligned}$$

where K contains all constants from the transforms and from the terms U_{LO}^2 and U_p and + or - correspond to frequencies $(\omega_{IF} + \omega_D)$ and $(\omega_{IF} - \omega_D)$ as shown in Figure 2. These points are color coded to indicate the radial speed. Equation (32) can be used to calculate the real and imaginary target signals at the point and the phase between them. If the two term in equation (32) are added, then

$$2KU_T(\bar{u}_i) \cos\theta(\bar{u}_i) = m_{+, \omega_D}(\bar{u}_i) + m_{-, \omega_D}(\bar{u}_i) \quad (33)$$

and if they are subtracted, then

$$2KjU_T(\bar{u}_i) \sin\theta(\bar{u}_i) = m_{-, \omega_D}(\bar{u}_i) - m_{+, \omega_D}(\bar{u}_i) \quad (34)$$

so

$$\tan\theta(\bar{u}_i) = \frac{j \left[m_{-, \omega_D}(\bar{u}_i) - m_{+, \omega_D}(\bar{u}_i) \right]}{\left[m_{+, \omega_D}(\bar{u}_i) - m_{-, \omega_D}(\bar{u}_i) \right]} \quad (35)$$

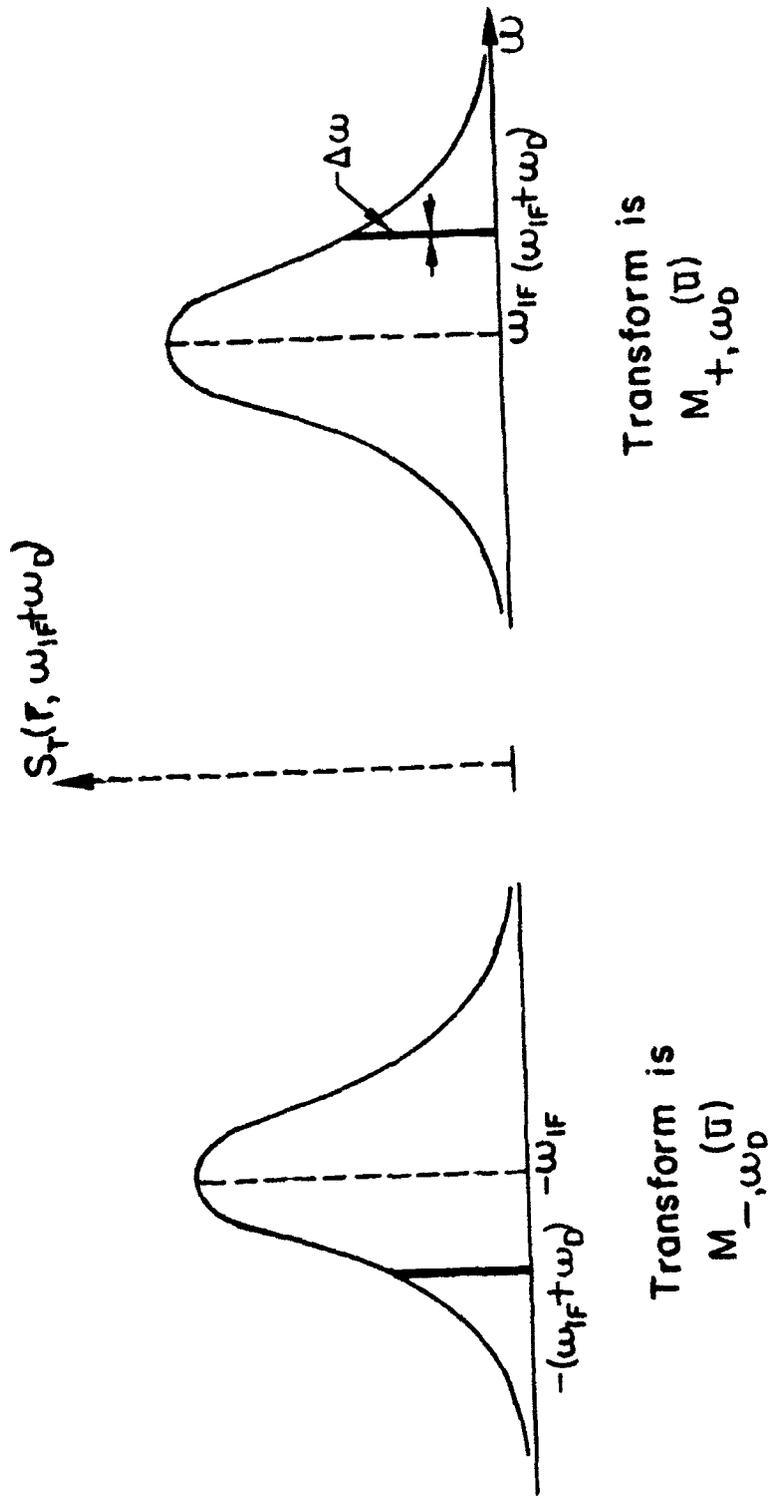


Figure 2 Meaning of Transformed Signals.

The speckled angle-angle-Doppler image is found by taking the product of the m 's and a quantity proportional to it is

$$\begin{aligned}
 I_s(\bar{u}_i) &\approx m_{-\omega_D}(\bar{u}_i) m_{-\omega_D}(\bar{u}_i) \\
 &= K^2 |U_T(\bar{u}_i)|^2 = k^2 I_s(\bar{u}_i)
 \end{aligned}
 \tag{36}$$

where I_s indicates the speckled intensity resulting from analyzing a single frame of data. If the target is allowed to rotate through a small angle between frames, then different speckle realizations can be averaged to approach the incoherent angle-angle-Doppler image. This averaging over a finite number of frames $j = 1, 2, \dots, N$ is

$$I_{\text{incoherent}}(\bar{u}_i) = \sum_{j=1}^N I_{sj}(\bar{u}_i)/N
 \tag{37}$$

As $N \rightarrow \infty$, the true incoherent image of the target in target space is evaluated

$$I_{\text{incoherent}}(\bar{u}_i) = \text{Limit}_{N \rightarrow \infty} \sum_{j=1}^N I_{sj}(\bar{u}_i)/N
 \tag{38}$$

CONCLUSIONS

In heterodyne detection one of the important parameters is the constant k_D . In our experiment where the IF is limited, it restricted our radial speed to be in the small of the order of $\mu\text{m/s}$ and for targets with extremely high speeds of the order of km/s the IF frequency would have to be in the GHz region. At both extremes angle-angle-Doppler imaging is difficult. Even at radial speeds of tens of m/s in both angle-angle imaging and in angle-angle -Doppler imaging spares detector arrays may be required. In all cases the motion of the target must be small in each frame and this possible can only be achieved by using sparse detector arrays with very possibly parallel processing of the data.

REFERENCES

1. S. E. Clark, L. R. Jones, and L. F. Desandre, Coherent Array Imaging, Appl. Opt. 30, 1804-1810 (1991)
2. S. E. Clark, private communication

3. W. Alexander, D. L. Clark, and C. A. Stewart, A CO₂ Laser Radar Configured for Doppler Imaging, Active Infrared Systems and Technology, Proc. SPIE 806, 102-109 (1987)
4. S. S. Eucker and M. L. Fleshner, Imaging Three-Frequency CO₂ Laser Radar, Laser Radar IV, Proc. SPIE 1103, 159-166 (1989)
5. J. L. Meyzonnette, B. Remy, and G. Saccomani, CO₂ Laser Imaging System, Coherent Laser Radar: Technology and Applications, Snowmass, Co. pp 251-254, July, 1991
6. A. Yariv, Optical Electronics, John Wiley and Sons, Fourth edition, 1991
7. J. H. Shapiro, B. A. Capron, and R. C. Harney, Imaging and Target Detection with a Heterodyne Reception Optical Radar, Appl. Opt. 20, 3292-3313 (1981)
8. A. L. Kachelmyer, Laser Radar Acquisition and Tracking, Laser Radar IV, Proc. SPIE 1103, 120-138 (1989)

A BRIEF STUDY OF PASSIVE VISCOUS DAMPING FOR THE SPICE BULKHEAD STRUCTURE

Thomas J. Thompson
Graduate Research Assistant
Department of Mechanical Engineering

Joseph R. Baumgarten
Professor
Department of Mechanical Engineering

Iowa State University
2024 H. M. Black Engineering Building
Ames, Iowa 50013

Final Report for:
Summer Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington DC

September 1992

A BRIEF STUDY OF PASSIVE VISCOUS DAMPING
FOR THE BULKHEAD STRUCTURE

Thomas J. Thompson
Graduate Research Assistant

Joseph R. Baumgarten
Professor

Department of Mechanical Engineering
Iowa State University

ABSTRACT

The SPICE Testbed at Phillips Laboratory is being used to evaluate the effects of structural vibration on line-of-site error for this strut built structure. A design incorporating active control and passive damping techniques is suggested to reduce the optical path distortion created in the vibrating structure. The passive viscous damping applied to the structure serves to aid the active control system stability in the cross-over and spill-over frequency range by producing a specified magnitude of damping in specified critical modes. This magnitude of damping is to be achieved by replacing the standard filament wound undamped struts with optimally placed D-struts which contain series and parallel combinations of springs and viscous dampers and produce damped vibration response from in-line strut deflection. This D-strut must replace standard struts in a tear-down of the bulkhead. The present study proposes to provide the requisite damping by adding on viscous damping at diagonal nodal locations in the bulkhead, circumventing the need to disassemble the SPICE bulkhead. The study shows specific increase in loss factor and improved damping ratio provided by the diagonal dampers when compared to in-line D-struts for specific modes and frequencies.

A BRIEF STUDY OF PASSIVE VISCOUS DAMPING
FOR THE BULKHEAD STRUCTURE

Thomas J. Thompson
Joseph R. Baumgarten

1. LINEAR ENERGY DISSIPATION

1.1 Introduction

Passive viscous damping is to be provided to the SPICE bulkhead by use of in-line D-struts or V-struts containing viscous damping elements [1]¹. In order to compare the energy dissipated per cycle and the attendant damping ratio ξ at a given mode of vibration, it is necessary to define loss factor η for a given design and develop the relationship between loss factor and damping ratio. Loss factor η is the basis for design of D-struts [2] and will be used here to compare energy dissipation capacity for structural cells containing in-line D-struts and proposed diagonal viscous dampers.

1.1.1 Kelvin Model of D-Strut

Figure 2.3 shows the simplified 3-element model of a D-strut where k_c and k_s represent the equivalent static and series stiffness, and c the effective viscous damping coefficient [1]. The dynamic response to an input force F is

$$c(\dot{x}_1 - \dot{x}_2) + k_c x_1 = F \quad (1.1)$$

$$-c(\dot{x}_1 - \dot{x}_2) + k_s x_2 = 0 \quad (1.2)$$

Taking Laplace transforms there results

$$cs(x_1 - x_2) + k_c x_1 = F(s) \quad (1.3)$$

$$cs(x_1 - x_2) - k_s x_2 = 0 \quad (1.4)$$

Solving Equation (1.4) for x_2 and substituting in Equation (1.3), one finds

$$\frac{F}{X_1}(s) = \frac{k_c k_s + (k_s + k_c) cs}{k_s + cs} \quad (1.5)$$

¹numbers in brackets cite REFERENCES

Substituting $s=j\omega$, defining the complex stiffness in terms of the real part of the stiffness and the loss factor η , gives

$$k(\omega) = k_r + jk_i = k_r(1 + j\eta(\omega)) = \frac{k_c k_s + (k_c + k_s) c(j\omega)}{k_s + c(j\omega)} \quad (1.6)$$

Clearing the complex fraction on the right yields

$$\eta(\omega) = \frac{k_s^2 c \omega}{k_c k_s^2 + (k_s + k_c) c^2 \omega^2} \quad (1.7)$$

Equation (1.7) can be manipulated to maximize η in a given design. This will be discussed in Section 2.

1.2 Damping Ratio of an Equivalent SDOF System

Figure 1.1 shows an equivalent single degree of freedom (SDOF) system representative of a single complex eigenvalue ω of a linearly damped complex structure. The complex frequency response to an external force $F(j\omega)$ is

$$(-\omega^2 m + k + j\omega c) X = F(j\omega) \quad (1.8)$$

Comparing Equations (1.7) and (1.8), for light damping with c less than one quarter of the critical value, dashpot impedance can be neglected [2], and

$$\omega_{eq} = \sqrt{\frac{k_r(\omega)}{m}} \quad (1.9)$$

The energy dissipated per cycle by the linearly viscous damped equivalent system of Figure 1.1 is found from

$$E_{dq} = \pi \xi_{dq} (2m\omega_{dq}^2) (x(t))^2 \quad (1.10)$$

The energy dissipated at coordinate x_1 of the D-strut of Figure 2.3 (the location of any lumped mass in FEM representation) is found from

$$E_d = \pi \eta k_r (x_1(t))^2 \quad (1.11)$$

Equating the energy dissipated, realizing the resonance condition $\omega = \omega_{eq}$ exists, gives

$$\xi_{dq} = \frac{\eta}{2} \quad (1.12)$$

where $\eta = \eta(\omega)$ is given by Equation (1.7). This establishes a basis for comparing loss factors for in-line and diagonally damped cells to arrive at optimal damping ratios.

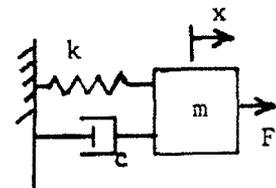


Fig. 1.1

2. COMPARISON OF IN-LINE DAMPING WITH DIAGONAL DAMPING

2.1 Analytical Development

2.1.1 Cell-Level Comparison

In order to compare in-line damping with diagonal damping, a fair basis of comparison had to be found. Since the diagonal damper is unable to carry static load like the in-line damper can, comparing loss factor in the basis of the single strut (or damper) would not be meaningful.

To maintain simplicity, a small group of struts which comprises a basic "building block," of the tetrahedron truss structure was chosen. This group of struts, or octahedral "cell" is shown in Figure 2.1. Comparison was made by placing a damping strut either across any two opposing nodes (diagonal damping), or in parallel with one or more of the existing struts, (in-line damping). The resulting loss factor for the cell gave a meaningful measure of damping. Loss factor can be defined as the ratio of energy dissipated to the energy stored as elastic strain energy.

2.1.2 Derivation of Loss Factor for Cell

In order to analytically determine loss factor for a cell, the following assumptions were made:

- base constrained against translation in any direction
- top node constrained in the x and y directions
- joints assumed "pinned," as in a truss
- small deformations
- massless struts
- force applied vertically at top joint

As with single struts, the cell loss factor was found by taking the ratio of the imaginary to real parts of the frequency-dependent complex stiffness.

$$\frac{F}{X}(\omega) = k_r + jk_i = k_r(1 + \eta j)$$

$$\eta(\omega) = \text{loss factor} = \frac{k_i}{k_r}(\omega)$$

$$\frac{F}{X}(\omega) = \text{force / displacement frequency response}$$

As before, ω is the rotational frequency, k_r is real stiffness, and k_i is imaginary stiffness.

2.1.3 Development

Using the above assumptions, spring-and-damper models of both configurations were formed. In effect, the octahedral structure without dampers became a spring in parallel with the spring-damper combination of the struts. As the spring rate of the structure was sought, absolute value of the axial force in all struts was found to be identical. This was taken to be true under dynamic loading, since the bars were assumed massless.

2.1.3.1 Diagonal-Strut

Figures 2.2 and 2.3 show the diagonal strut octahedral configuration and the corresponding spring-damper model. Stiffness of the structure was found by determining the displacement at the top node due to a given force. A downward vertical force P at the top node of the structure would result in an axial force (tensile or compressive) on each strut of

$$F = 0.354P$$

The resulting axial deformation in each strut is then

$$d_1 = \frac{Fl}{AE}$$

where d_1 is deformation, l is strut length, A is strut diagonal-sectional area, and E is its elastic modulus. The struts can be grouped into three sets: the top four, the horizontal four, and the bottom four. Each set contributes equally to top-node displacement, d_h .

$$d_h = \sqrt{2}d_{top} + \sqrt{2}d_{bot} + \sqrt{2}d_{horiz} \quad (2.1)$$

$$d_h = 3\sqrt{2} \frac{Fl}{AE} = 3\sqrt{2} 0.354 \frac{Pl}{AE}$$

$$k_c = \frac{P}{d_h} = \frac{PAE}{3\sqrt{2} 0.354 Pl} = \frac{AE}{3\sqrt{2} 0.354 l} \quad (2.2)$$

The loss factor for the combined diagonal strut octahedral cell is derived identically to that for a single D-strut of Eqn. (1.7). Here, subscript c refers to the combined static stiffness of the cell. Thus

$$\eta(\omega) = \frac{k_s^2 c \omega}{k_c k_s^2 + (k_s + k_c) c^2 \omega^2} \quad (2.3)$$

To find the frequency at which maximum loss occurs, the derivative of the above expression with respect to ω , is set to zero. The result is

$$\omega_{\max} = \frac{k_s}{c} \sqrt{\frac{k_c}{k_s + k_c}} \quad (2.4)$$

Substituting back into Equation (2.3) results in

$$\eta_{\max} = \frac{k_s}{2k_c} \sqrt{\frac{k_c}{k_s + k_c}} = \frac{K}{2} \sqrt{\frac{1}{K+1}} \quad (2.5)$$

and from Equation (2.4)

$$c(\omega) = \frac{k_c^2}{\omega_{\max} \sqrt{k_s(k_c + k_s)}} = \frac{k_c}{\omega_{\max} \sqrt{K(1+K)}} \quad (2.6)$$

where

$$K = \frac{k_s}{k_c} = \frac{\frac{E_d A_d}{\sqrt{2} l}}{\frac{E_s A_s}{1.06 \sqrt{2} l}} = 1.06 \frac{E_d A_d}{E_s A_s}$$

and the s subscript in this equation refers to properties of the outer struts, the c subscript refers to the cell, and the d subscript refers to properties of the diagonal strut.

Thus the achievable loss factor depends not on c, but on the ratio of the effective spring rate of the structure to the spring rate of the strut in series with the damper. This is also the case for a single in-line strut [1]. The frequency at which this maximum loss factor occurs can be set by specifying c. Table 1 gives η_{\max} for various values of K.

2.1.3.2 Four In-line Struts

Figure 2.4 shows the octahedral arrangement for assessing the effectiveness of in-line damped struts. To maintain symmetry, the top four struts were replaced by in-line damped struts. It was assumed that each of these four struts consisted of the spring-damper shown in Figure 2.5. The stiffness of the outer, load-bearing shell is k_A , that of the inner element transmitting force to the damper is k_B , and the damping coefficient is c. As in the diagonal-damping case, all struts will carry a load of magnitude F. As shown in Equation (2.1), each of the three subsets of struts within the octahedron will have its own contribution to the displacement d_h of the upper node. Therefore, the octahedron can be modeled as shown in Figure 2.6.

$$d_h = \sqrt{2} (d_{top} + d_{bot} + d_{diag})$$

$$d_{top} = \frac{k_B + c s}{k_B c s + k_A k_B + k_A c s} F$$

$$d_{top} = d_{bot} = \frac{F}{k_A} = \frac{l F}{A E}$$

where k_A is the stiffness of one of the struts, also equal to the stiffness of the outer shell of the in-line, damped strut. Rearranging results in the transfer function

$$\frac{P}{d_h}(s) = \frac{F}{0.354 d_h} = \frac{2.82}{\sqrt{2}} \left(\frac{1}{\frac{1}{k_A} + \frac{1}{k_B} + \frac{k_B + c s}{k_B c s + k_A k_B + k_A c s}} \right)$$

Substituting $s = j\omega$ to obtain a frequency response function and setting

$$\frac{P}{d_h}(\omega) = k_r (1 + \eta(\omega) j)$$

results in

$$\eta(\omega) = \frac{\omega c k_A k_B^2}{3 k_A^2 k_B^2 + \omega^2 c^2 (k_A + k_B) (3 k_A + 2 k_B)}$$

As in the diagonally-damped case, η is maximized with respect to frequency. The result is

$$\omega_{max}^2 = k_A^2 \frac{3 K_1^2}{(1 + K_1) (3 + 2 K_1) c^2}$$

$$\eta_{max}(\omega) = \frac{\sqrt{3}}{6} K_1 \sqrt{\frac{1}{(1 + K_1) (3 + 2 K_1)}}$$

where

$$K_1 = \frac{k_B}{k_A} = \frac{E_B A_B}{E_A A_A}$$

Again, it is seen that maximum loss factor is dependent upon the stiffness ratio, but not upon c . The frequency at which the maximum loss factor occurs can be specified by varying c .

Table I also shows loss factor η_{max} for various values of K_1 . Values of K and K_1 in the same row indicate the diagonal-strut body would have cross-section equal to that of the inner struts in the in-line case. It is apparent that for this condition and loading, a diagonally-damped octahedron achieves a much higher overall loss factor.

It can be further shown that the achievable loss factor for an octahedron in which all struts contain equal in-line damping cannot exceed the loss factor for any one strut tested separately.

2.2 Computer Simulation Results

2.2.1 Model Description

In order to confirm the results obtained analytically, a model of the octahedron was simulated using the finite element software MSC/NASTRAN.

The model is shown in Figure 2.1. It consists of 12 struts, each containing one element. The properties of the bar elements were specified to emulate the carbon-epoxy tubes used as struts on the SPICE structure. These properties are given in Table II.

The two cases analyzed in the previous section were created by adding additional bar elements in series with CVISC (pure damping) elements in the appropriate locations. The damping strut bar elements spanned all but one inch of the length of each damped strut. The series damper accounted for the other inch, and the intervening node was constrained to stay on the axis of the strut. The bottom node was constrained against all translation, and against "spinning" rotation. The top node was constrained from translating horizontally, and any force was applied vertically to this node.

Table I. Maximum Loss Factor for Values of K and K_1

Diagonal Damping		In-line Damping	
K	η_{max}	K_1	η_{max}
1.06	0.369	1.0	0.091
2.12	0.600	2.0	0.126
3.18	0.777	3.0	0.144
4.24	0.962	4.0	0.156
5.30	1.056	5.0	0.163
106.0	5.120	100.0	0.202

Table II. Strut Properties for the SPICE Bulkhead

Property	Value
Length	1.03 meters
E modulus	1.1E+11 N/m ²
G modulus	2.65E+9 N/m ²
Area	2.55E-4 m ²
Density	8024 kg/m ³
I_{xx}, I_{yy}	5.0E-8 m ⁴
I torsion	1.0E-7 m ⁴

Both frequency response and modal dynamic analysis were performed on the NASTRAN model. The forced frequency response was used to confirm loss factor values, and to picture the effect of damping. The complex modal analysis was used to compare the effectiveness of the two damping strategies in achieving high modal damping.

2.2.2 Forced Response

In order to confirm the accuracy of the analytical loss factor values, forced responses were run on the diagonally-damped and in-line damped octahedrons using NASTRAN. With a sinusoidal vertical force exerted on the top node of the model over a range of frequencies, NASTRAN computed the resultant complex displacement (real and imaginary parts, or magnitude and phase).

As in the case for experimental analysis of single in-line damping elements, loss factor was measured as the imaginary divided by the real components of the vertical displacement of the top node

$$\eta = \frac{d_i}{d_r} = \tan\phi$$

This method is accurate if mass effects are negligible, hence the value of c was set using Equation(2.6) so that the maximum value of η fell in a frequency region where there are no eigenvalues. For this octahedron, c was set so η would peak around 150 Hz. Values of cross-sectional area for the diagonal strut and inner struts in the two cases were set equal, so that $K=5.3$, and $K_1=5.0$.

2.2.2.1 Results for Diagonal Damping

When the forced response was performed on the vertical diagonally-damped octahedron, the displacement plot in Figure 2.7 resulted. Analytical values of ϕ were also plotted using the relation

$$\phi = \tan^{-1}\eta$$

This figure shows that η peaks out around 150 Hz, reaching a value of -48° . This corresponds to a loss factor of 1.11, confirming the prediction in Table I made by analytical means.

2.2.2.2 Results for In-Line Damping

Figure 2.8 shows the forced response of the octahedron with four in-line damped struts. The maximum value of ϕ reached -8.22° , corresponding to a loss factor η of 0.145. This is also very close to the analytical result. These forced response cases confirm that a much higher loss factor can be attained

using diagonal damping for this case of dynamic loading. This result also matches the analytical prediction of Table I.

2.2.2.3 Other Configurations

To test the effectiveness of damping configurations under different loading conditions, the forced responses in Figures 2.9 and 2.10 were run. Figure 2.9 shows the magnitude and phase of the response in the case where a diagonal damper was oriented horizontally across two opposite nodes of the octahedron. With the same function, a much lower loss factor of 0.0499 corresponding to a phase angle of -2.86° was realized.

Figure 2.10 shows the response in the case where the four in-line damped struts were oriented so they all met at one of the "outer" or "side" nodes of the octahedron. The phase angle in this case is -10° , giving a loss factor of 0.17. Thus, the four in-line strut configuration actually had a better loss factor under this loading case. It appears that this configuration with a relatively large number of dampers is less sensitive to loading orientation than the diagonally-damped octahedron.

The two diagonally-damped cases presented may be seen as best and worst loading cases. Since this type of damping is more sensitive to orientation, though very effective, the issue of optimal placement of diagonal-dampers in large space structures is very important.

2.2.3. Eigenvalue Analysis

In a large space structure, such as SPICE, damping is applied in order to control vibration in resonant modes. Therefore, eigen analysis was performed on MSC/NASTRAN to assess the effectiveness of diagonal-damping in a structure. Modes were found on NASTRAN using the inverse power method to obtain complex eigenvalues. Modal damping coefficients computed by the program were according to the classical definition

$$\xi = \frac{s_i}{\sqrt{s_r^2 + s_i^2}}$$

where s_r and s_i are the real and imaginary components of the complex eigenvalue.

2.2.3.1. Description of Structure Analyzed

The structure which was analyzed for the effect of diagonal struts on modal damping was the bulkhead platform of the SPICE structure. It is shown in Figure 2.11. This structure was chosen because it is typical of the large space structures which need damping, because it consists of a series of octahedral strut

cells, and because there already existed a finite element model which had been correlated which modal test data.

Since the structure is large, it contains many locations at which damping struts could be inserted. Therefore, due to time constraints, the locations of struts had to be determined by intuition. The goal was to assess the damping potential of diagonal damping as compared to that of in-line damping.

2.2.3.2. Undamped Case

The results of the eigen analysis on the undamped structure are given in Table III, and mode shape (using PAT-RAN P/FEA) corresponding to the first six eigenvalues are shown in Figure 2.12. The undamped results are presented so that mode shapes and the effect of various damping on modal frequencies can be seen.

Table III. Eigenvalues of Undamped Truss Structure

Root	Frequency, Hz
1	54.36
2	54.36
3	82.37
4	87.83
5	106.51
6	119.38

2.2.3.3. Diagonally-Damped Structure

The modes for the case of diagonal damping on the structure are shown in Table IV, along with modal damping. For this run, the damper was on one of the outside corners of the structure, between nodes 51 and 52 (see Figure 2.11). This location was chosen because modal undamped displacement diagrams showed a large amount of shape distortion for this area in certain modes. The value of c was set at 41,100 N-s/m so that the maximum loss factor (for an octahedron) would occur at 150 Hz.

As Table IV shows, for one damped strut, good modal damping was achieved. The damping coefficient was better than 0.01 for modes 2 and 6, and modes 3, 4, and 5 achieved coefficients better than 0.001.

2.2.3.4. Eigen Analysis of In-line Damped Structure

The modes for the structure containing in-line damping struts are shown in Table V. The in-line damping strut was inserted into the NASTRAN model between nodes 28 and 34. This location was chosen because this strut appeared to carry some of the highest strain energy in various modes among the struts which were

evaluated. This may have been due to being on the top deck of struts (outer "filament").

The damping coefficients for this case were lower than in the diagonally-damped case. Modes 2 and 4 achieved damping coefficients above 0.005. Modes 3, 5, and 6 were in the range of 0.0005 to 0.001.

Table IV. Modes and Damping for Diagonally-Damped Truss

Mode	Freq. Hz	Damping Coef.
1	53.2	1.6E-4
2	55.0	1.26E-2
3	82.1	2.42E-3
4	88.4	3.01E-3
5	107.	7.6E-3
6	125.	1.09E-2

2.2.4 Discussion of Results

2.2.4.1. Loss Factor

The excellent correlation between loss factor obtained analytically and computationally for octahedral structures confirms that diagonal-damping could potentially provide much higher loss factor for certain load cases, even when a smaller number of dampers is used. Forced responses of various damping orientations show sensitivity to load orientation and highlight the need for optimal strut placement.

2.2.4.2. Eigenvalue Analysis

The eigenvalue analysis for diagonal and in-line damping cases both involving one damped strut in a large truss implies that diagonal damping achieves modal damping values comparable to those of in-line damping. In the case given, diagonal performance was actually better; however, damper locations for this run were chosen by intuition, and do not necessarily represent optimal positions. The test did confirm that significant damping can be achieved in a large tetrahedron truss structure by means of diagonal damping.

2.2.5. Research Opportunities

The above results open several opportunities to be researched. Cross diagonal damping has shown potential of providing much higher loss factors using fewer elements than in-line strut damping. This could result in lighter structures.

Since diagonal dampers could be installed without disassembling the original structure, they could be used for selective damping after structural assembly. Indeed, on the SPICE structure, most node balls already have threaded holes in locations where diagonal damping could be added. This feature could also allow

Table V. Modes and Damping for In-line Damped Truss

accurate finite element-modal analysis correlation on an undamped structure, with necessary damping added to the structure without disassembling and disrupting the structure.

Since diagonal dampers would not have to carry static loads, it is possible that one could be built using a simpler and lighter design than existing in-line damping struts (D-struts and V-struts) [1].

Mode	Freq. Hz	Damping Coef.
1	54.3	5.0E-6
2	54.3	7.7E-3
3	82.4	8.8E-4
4	87.4	5.6E-3
5	106.	6.8E-4
6	119.	9.8E-4

The directional nature of the diagonal damping struts shows the need of research into optimal strut placement, since some orientations of the strut appear to be far more effective than others. Since the frequency at which maximum loss factor can occur is dependent upon c , optimal "c" damping coefficients for each diagonal strut should be found to truly optimize damping. The trade-off between values of damper series stiffness (resulting in higher loss factor) and structural weight could also be investigated.

Finally, results should be confirmed experimentally. The octahedral loss factor results should be reproduced using physical diagonally-damped struts. Then the SPICE structure could provide an extremely good bed for testing the effectiveness of placing diagonally-damped struts in various, even optimized locations, since its modal properties have already been accurately determined.

REFERENCES

- [1] Yiu, Y.C., Davis, L.P., and Kienholz, D.A., "Development of High Frequency Passively Damped Struts, SPICE Program", Final Report for subtask 02-08, Task 4, April, 1992, pp. 21-23, 35-36.
- [2] Yiu, Y.C., and Ginter, S.D., "Dynamics of A Class of Viscously Damped Struts", Proceedings, Damping'91, USAF Wright Lab., Flight Dynamics Dir., WL-TR-91-3078, Vol.1, Feb. 13-15, 1991, San Diego, Cal., pp. ccc1-ccc13

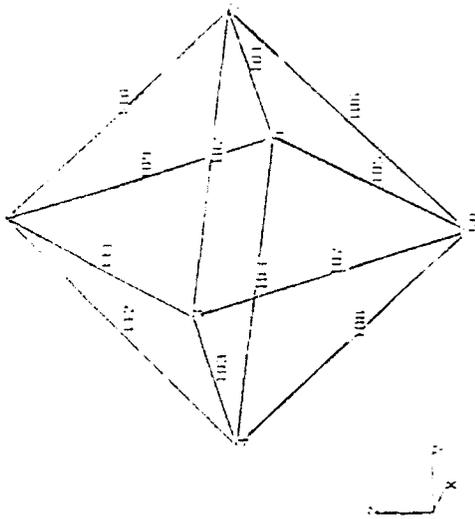


Figure 2.1 Octahedral Cell

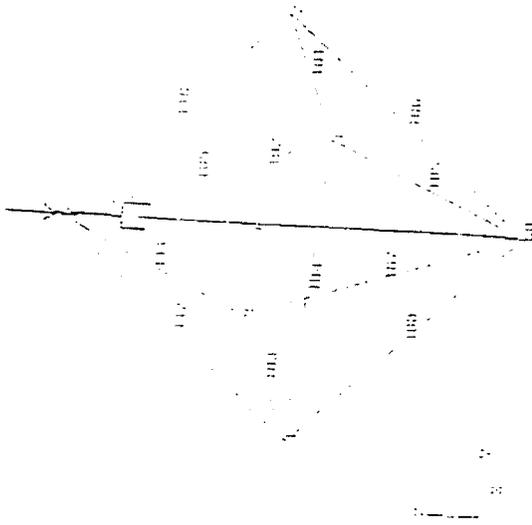


Figure 2.2 Octahedral Cell with Diagonal Strut

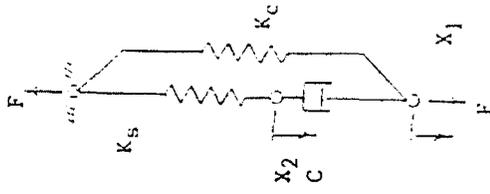


Figure 2.3 Diagonal Strut and Cell Equivalent Model

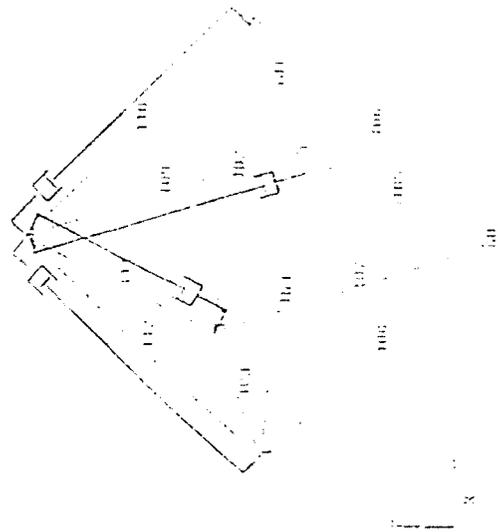


Figure 2.4 Octahedron with Four In-line Struts

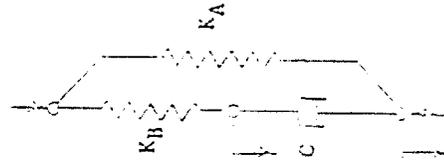


Figure 2.5 Model of In-line Damped Strut

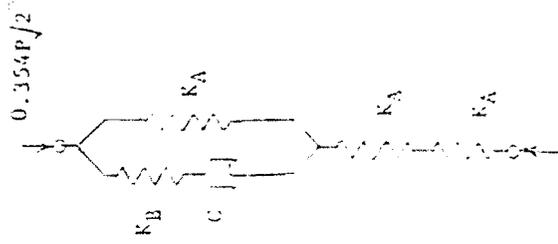


Figure 2.6 Equivalent Model of Octahedron with Four In-line Dampers

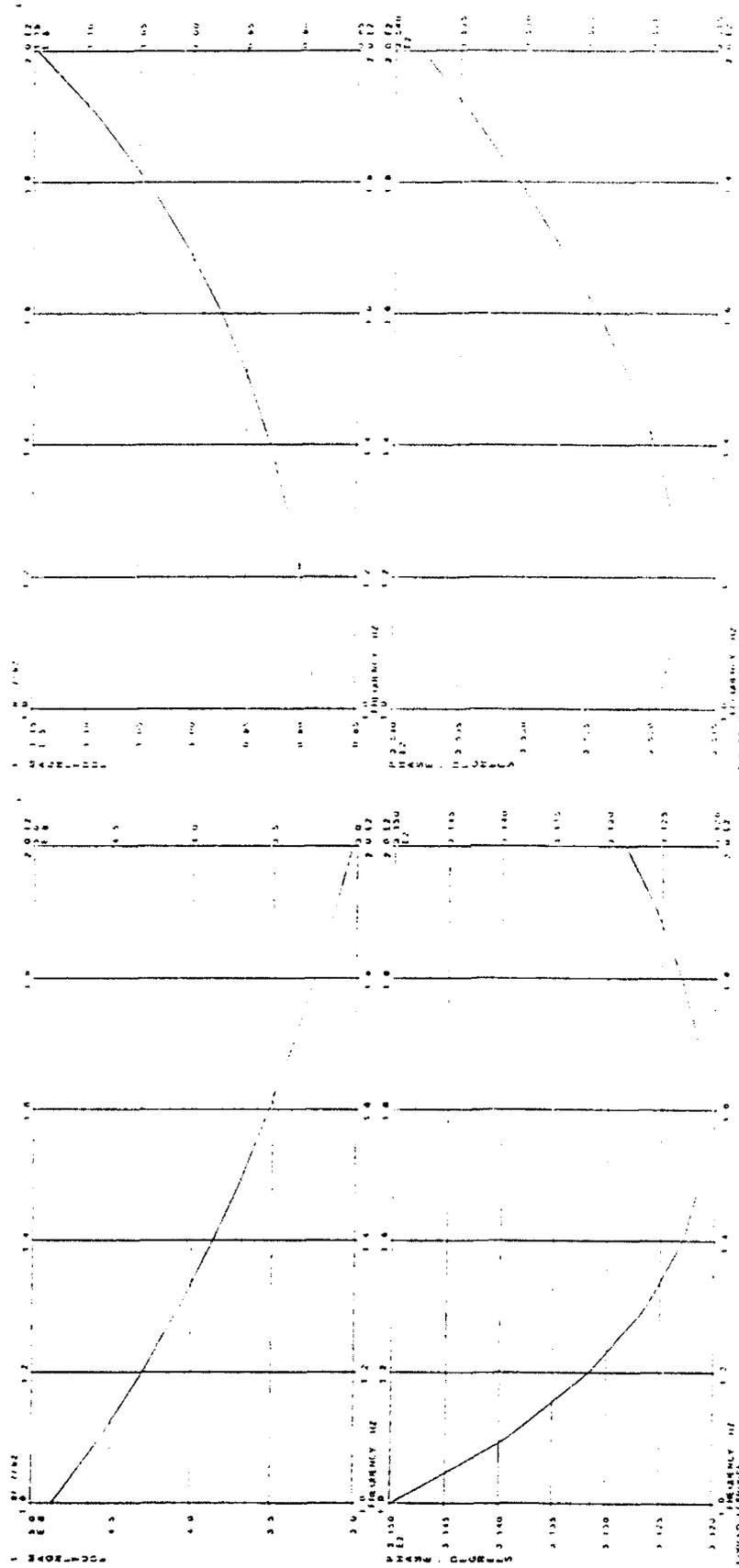


Figure 2.7 Forced Displacement Response For A Diagonally Damped Octahedral Cell

Figure 2.8 Forced Displacement Response For Four In-line Damped Struts in Cell

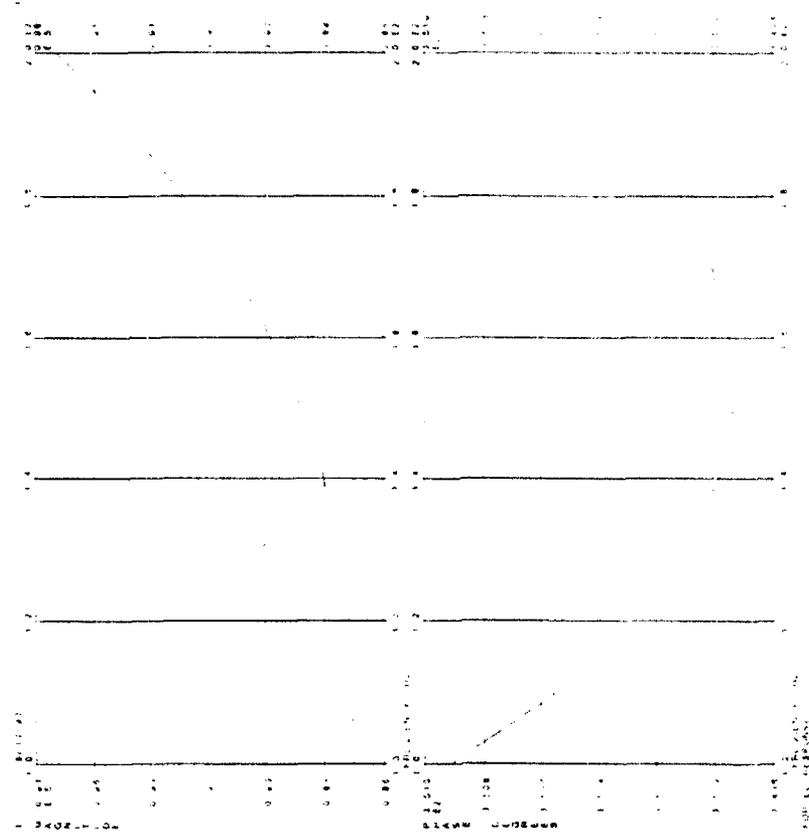


Figure 2.9 Forced Displacement Response
For Horizontal Diagonally Damped
Cell

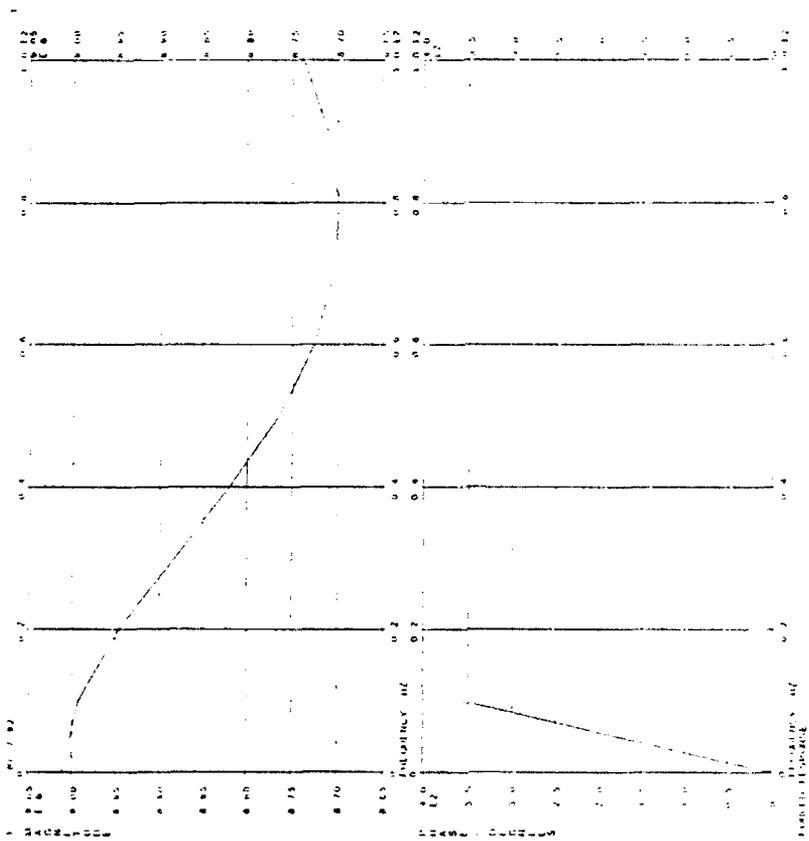
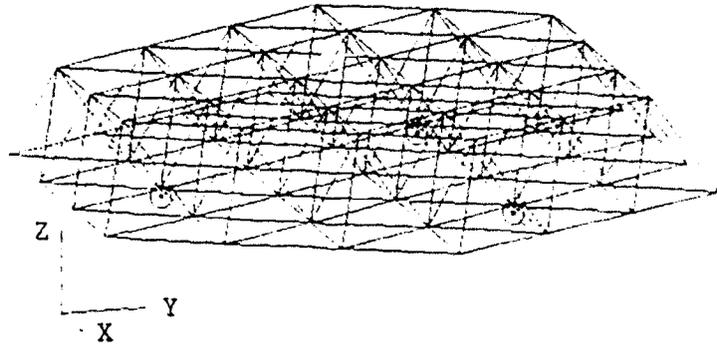
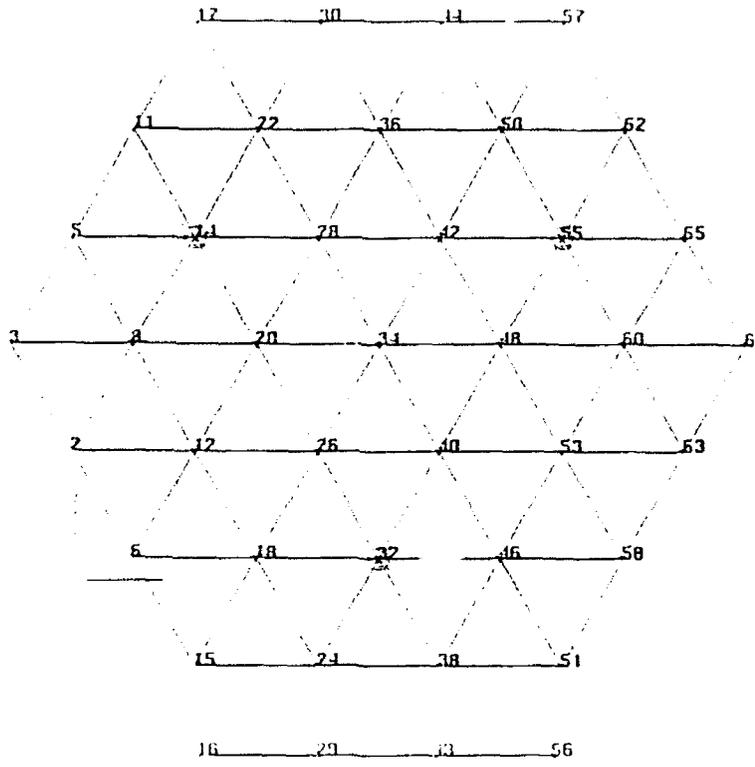


Figure 2.10 Forced Displacement Response For
Horizontal In-line Damped Cell

Oblique View



Lower Deck



Upper Deck

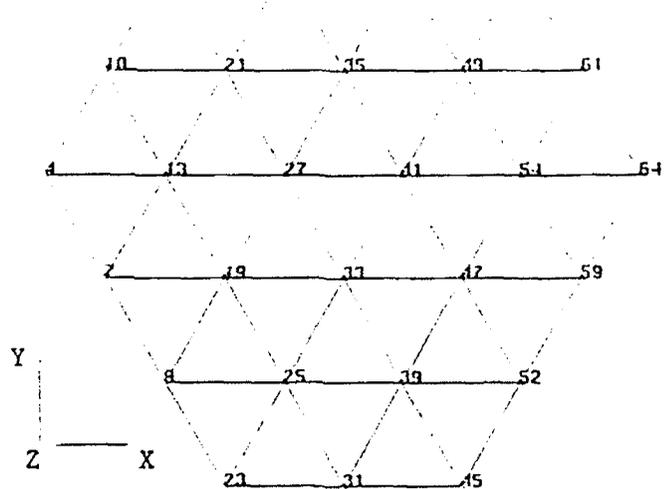
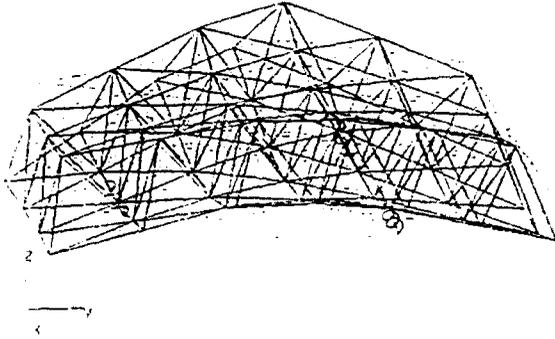
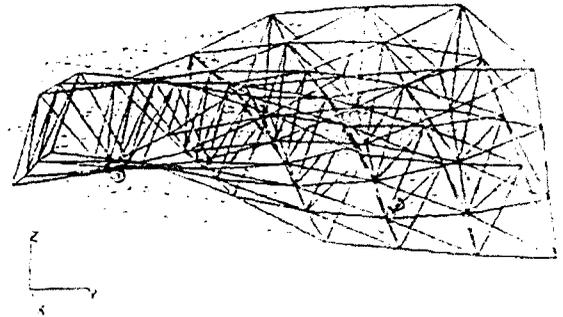


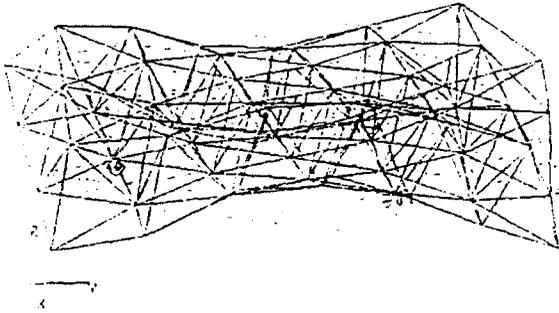
Figure 2.11. SPICE Bulkhead Truss



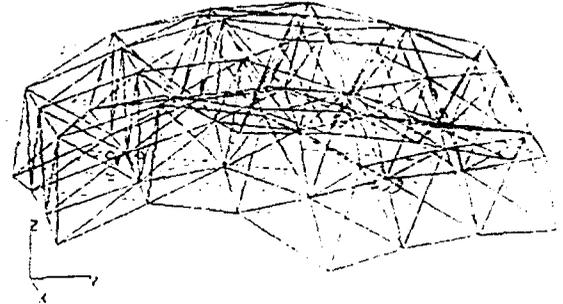
MODE = 1 FREQUENCY = 04-aug-92 09:33:21 p/lee 2.5-1
35.0 MAX DEF=0.10318



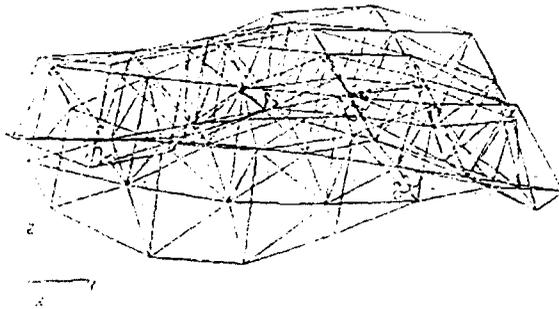
MODE = 2 FREQUENCY = 04-aug-92 09:33:21 p/lee 2.5-1
45.0 MAX DEF=1.025



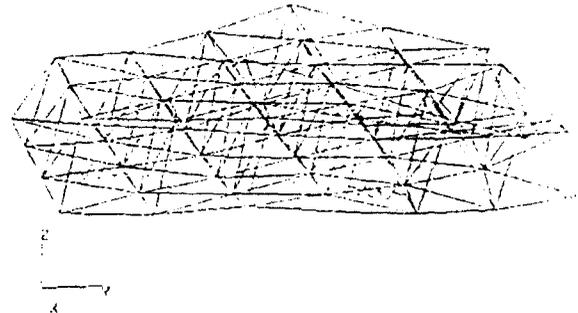
MODE = 3 FREQUENCY = 04-aug-92 09:33:21 p/lee 2.5-1
31.1 MAX DEF=0.02017



MODE = 4 FREQUENCY = 04-aug-92 09:33:21 p/lee 2.5-1
29.6 MAX DEF=1.025



MODE = 5 FREQUENCY = 04-aug-92 09:33:21 p/lee 2.5-1
49.9 MAX DEF=0.05424



MODE = 6 FREQUENCY = 04-aug-92 09:33:21 p/lee 2.5-1
131.1 MAX DEF=0.05

Figure 2.12 First Six Modes of SPICE Bulkhead

**CALIBRATION TECHNIQUES FOR A
LOW ENERGY X-RAY IRRADIATION CHAMBER**

**Raymond D. Bellem
Associate Professor
Department of Electrical Engineering**

**Embry-Riddle Aeronautical University
3200 Willow Creek Rd.
Prescott, AZ 86301**

**Final Report for:
Summer Research Program
Phillips Laboratory**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.**

September 1992

CALIBRATION TECHNIQUES FOR A LOW ENERGY X-RAY IRRADIATION CHAMBER

Raymond D. Bellem
Associate Professor
Department of Electrical Engineering
Embry-Riddle Aeronautical University

Abstract

Analytical and experimental techniques required to calibrate the radiation dosimetry of a new low energy (8-160 keV) X-ray facility were developed and demonstrated at the Phillips Laboratory (PL). Theoretical calculations and experimental measurements from NIST calibrated electronic dosimeters, correlate within 20 percent over the range 8-50 keV. This correlation makes the PL X-ray cell operational up to 50 keV. The facility is capable of continuous (and uniform) dose-rate levels up to 200 Krad(Si)/min. over a target, and is a viable adjunct to Cobalt-60 gamma irradiations for total dose measurements.

CALIBRATION TECHNIQUES FOR A LOW ENERGY X-RAY IRRADIATION CHAMBER

Raymond D. Bellem

INTRODUCTION

Evaluating microelectronic parts and subsystems to the high total dose radiation environments required on some space and missile systems dictate the use of expensive simulation tools, such as a high-dose rate Cobalt-60 gamma irradiator or a continuous pulsed Linear Electron Accelerator (LINAC). To reduce the high screening cost, many device manufacturers and end users are evaluating and qualifying their microelectronic space parts on low energy X-ray test systems such as the Aracor Model 4100 Automatic Semiconductor Irradiation System. This system generates 8-50 keV X-rays and is designed to perform high total dose irradiations and production screening of semiconductor devices at the wafer stage of fabrication {1}. However, this test system is limited to irradiating chip areas of only a few centimeters in diameter. At the Phillips Laboratory, the need exists to evaluate larger structures and to study the interaction between microelectronic packaged devices, boards and small subsystems when simultaneously exposed to ionizing radiation environments. To meet this need, personnel at Phillips designed and built a large low-energy radiation test cell which will accommodate electronic assemblies up to one meter in diameter. A Philips Model MCN 164 X-ray Unit provides the radiation source whose output spectrum is "tunable" for energies in the range 8 to 160 keV. Following the ASTM standard guide for using X-ray testers on microelectronic devices being drafted by Brown at NRL [2], procedures used to calibrate such a facility are addressed in this paper.

APPROACH

The characterization and calibration of the radiation dosimetry of a large X-ray facility with a continuous energy spectrum is a complex problem requiring extensive analytical and experimental measurements. On the analytical side, electron-photon transport codes are required to calculate the X-ray tube spectrum and the radiation transport/deposition in the structures being irradiated. Experimental measurements with calibrated X-ray detectors are also required to validate the theoretical calculations.

As a first step, the output intensity spectrum of the PL's X-ray tube was calculated by Dozier at the Naval Research Laboratory using electron transport techniques developed at NRL [3]. A composite of the data is plotted in Fig. 1

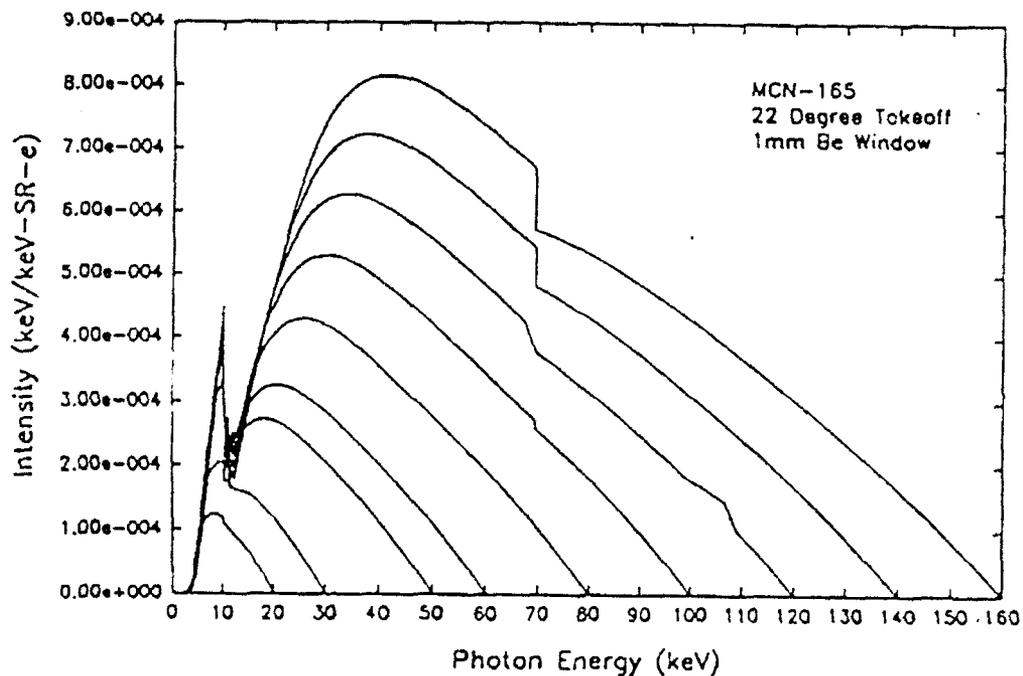
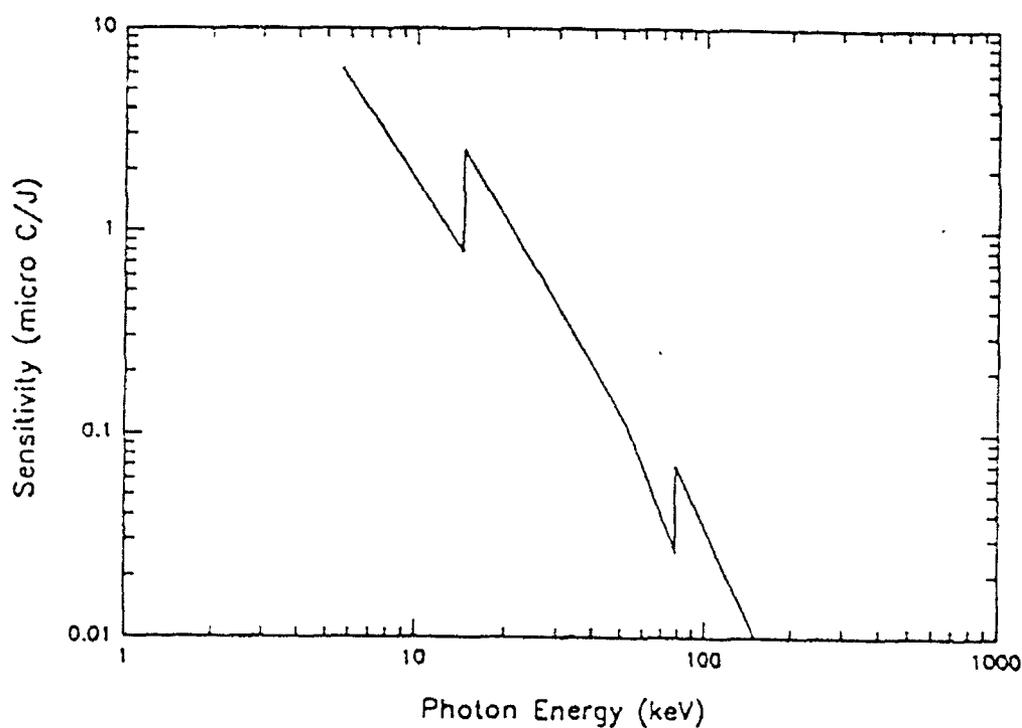


Figure 1: Calculated X-ray Tube Spectra

The radiation shielding, transport and deposition were then calculated using two other electron-photon transport computer codes called "PHOTCOEF" and "ITS". PHOTCOEF uses a semi-empirical model developed at the Rome Laboratories [4] and is commercially available for use on personal computers. Also, a computer program was developed by the author [5] to integrate the incident spectral intensity data shown plotted in Fig. 1 with detector sensitivity data shown plotted in Fig. 2, resulting in calculated diode currents. Diode current responses were calculated for the diode dosimeters irradiated at different tube operating conditions and at numerous test locations within the X-ray test cell.



Complementing these calculations were experimental measurements using three types of radiation detectors. The detectors selected were: a SEMIRAD (Secondary Electrons Mixed Radiation Dosimeter) vacuum diode, two PIN diodes and 75 PMOS FETS. The vacuum diode and the PIN diodes were calibrated at the SAIC-DNA X-ray calibration facility located in Albuquerque. This facility, traceable to NIST, uses bremsstrahlung-induced K-fluorescent monoenergetic X-rays of known intensities in the 1 to 110 keV energy range to calibrate the diode sensitivity.

The SEMIRAD is an evacuated chamber constructed to collect secondary emissions from a gold film cathode within the chamber. This vacuum diode, calibrated to energies up to 160 keV [6] is used regularly to measure "blackbody-like" spectral intensities in underground nuclear test at the Nevada Test Site. In this effort, the diode is used to measure the ionizing radiation field intensity over the full operating range of the irradiator within the PL's X-ray cell.

The two PIN diodes, commercially available from Aracor, are calibrated to directly measure dose rate in silicon [1]. The present calibration of these diodes is limited to X-ray energies up to 60 keV maximum. These diodes were used to map the absorbed dose in rad(Si) at various locations within the cell for energies up to 50 keV. Diode calibration techniques and results at higher energy X-rays using the SAIC-DNA facility, will be reported in a later paper.

The PMOS transistors were used to perform correlation studies between a calibrated Cobalt-60 gamma source and the low-energy X-ray source. These transistors, provided by the Sandia National Laboratory, use a state-of-the-art 180 angstrom radiation hardened gate oxide process. Using threshold voltage shift vs. total dose for these devices, and taking into account corrections for dose enhancement and electron-hole recombination per Table I of the ASTM guidelines [2], comparisons between Cobalt-60 and X-ray irradiations were made.

The measured data gathered from the above group of sensors not only generates accurate dosimetry within the X-ray chamber, but also provides confidence in the use of analytical calculations for test analysis.

EXPERIMENTAL MEASUREMENTS

SEMIRAD MEASUREMENTS. The experimental set-up for measuring the ionizing radiation field intensities with the SEMIRAD diode is illustrated in Fig. 3.

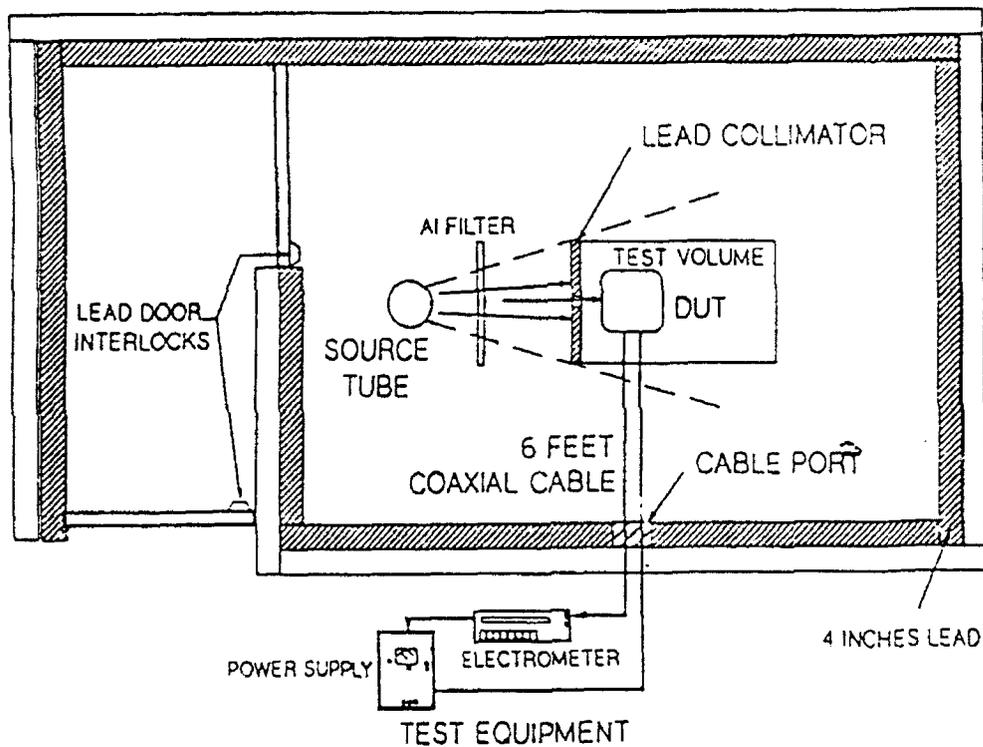


Fig. 3: PL X-ray cell (8-160 keV)

Shown is a Device-Under-Test (the diode) being irradiated by the X-ray source tube through an aluminum filter and lead collimator. The aluminum filter was placed between the source and the diode to insure that X-ray energies below 4 keV are filtered out in accordance with the ASTM

Guidelines. The collimator was used to limit the region irradiated. A thin piece of mylar was placed between the lead collimator and the diode to insure low energy scattered electrons do not reach the Beryllium window of the diode. To insure the collection of all the secondary electrons generated within the diode, a DC bias of -30 volts was employed across the diode. A Keithley Electrometer was used to measure the resulting steady state current generated by the irradiated diode. Current readings were recorded at different locations to map the radiation intensity in the X-ray cell. Table I, column 6 shows examples of the measured current using this technique.

Table I

Anode Voltage (KV)	Tube Current (mA)	Distance (cm.)	Intensity (keV/cm ² /sec)	Calculated Current (pA)	Measured Current (pA)	Error %
20	25	25	4.3 E11	375	419	+12
20	25	50	1.07 E11	143	82.5	+7.5
20	25	75	4.77 E10	30.2	28.7	+4.9
30	25	25	9.78 E11	919	913	-0.7
30	25	50	2.44 E11	193	198	+2.3
30	25	75	1.09 E11	73.6	78.0	+5.6
50	25	25	2.53 E12	1956	1605	-17.9
50	25	50	6.33 E11	419	348	-16.9
50	25	75	2.81 E11	165	137	-17.5

PIN DIODE MEASUREMENTS. With the exception of the diode bias voltage and the type of collimator employed, the experimental set-up for the PIN diode is very similar to that for the SEMIRAD. As discussed by Palkuti [1], the operating bias voltage of the diode must be zero volts. The diode radiation shield and collimator aperture assembly is constructed of

tantalum. Since the present diodes are not operated above 60 keV no further shielding is required. An example plot of X-ray tube current (intensity) vs. dose rate is shown in Fig. 4.

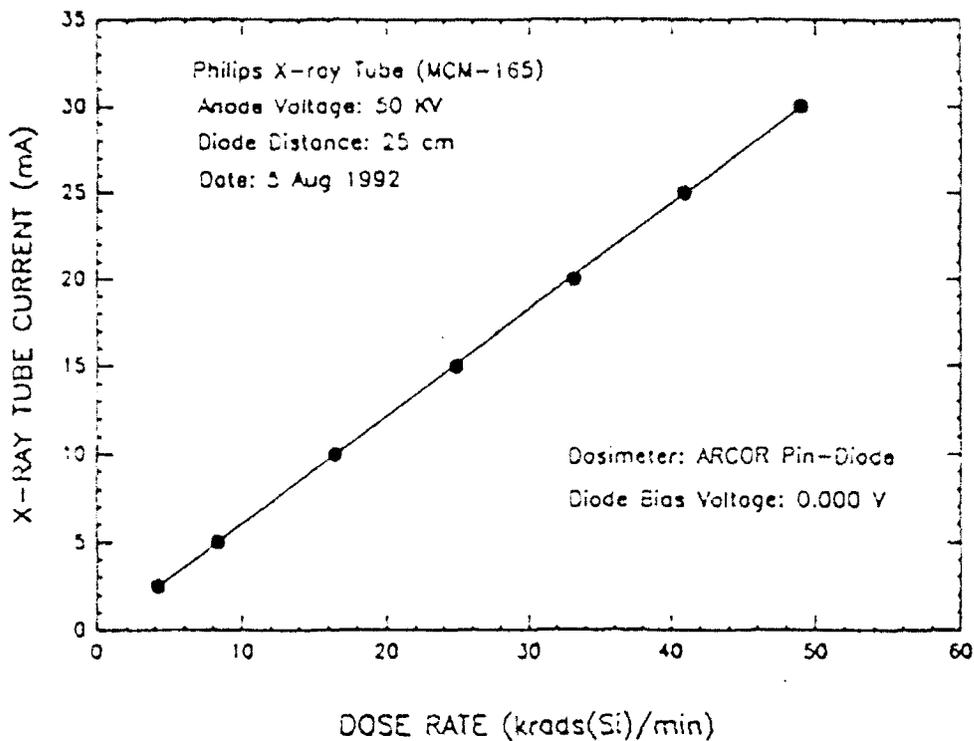


Figure 4: Facility Dosimetry Plot Using PIN Diode

PMOS TRANSISTOR MEASUREMENTS. Delidded PMOS devices were first irradiated in the calibrated PL free-field Cobalt-60 gamma facility, and gate threshold voltage shifts versus total dose were recorded "in situ". Log-log plots of the change in gate threshold voltage versus applied total dose for three dose-rate levels (6, 12, and 20 Krad(Si)/min.) are presented in Fig. 5.

The same class of PMOS transistors were then exposed at a dose rate level of 16.4 Krads(Si)/min. (as calibrated by the PIN diode) in the X-ray environment using a 50keV spectrum. Representative data is plotted in Fig. 6.

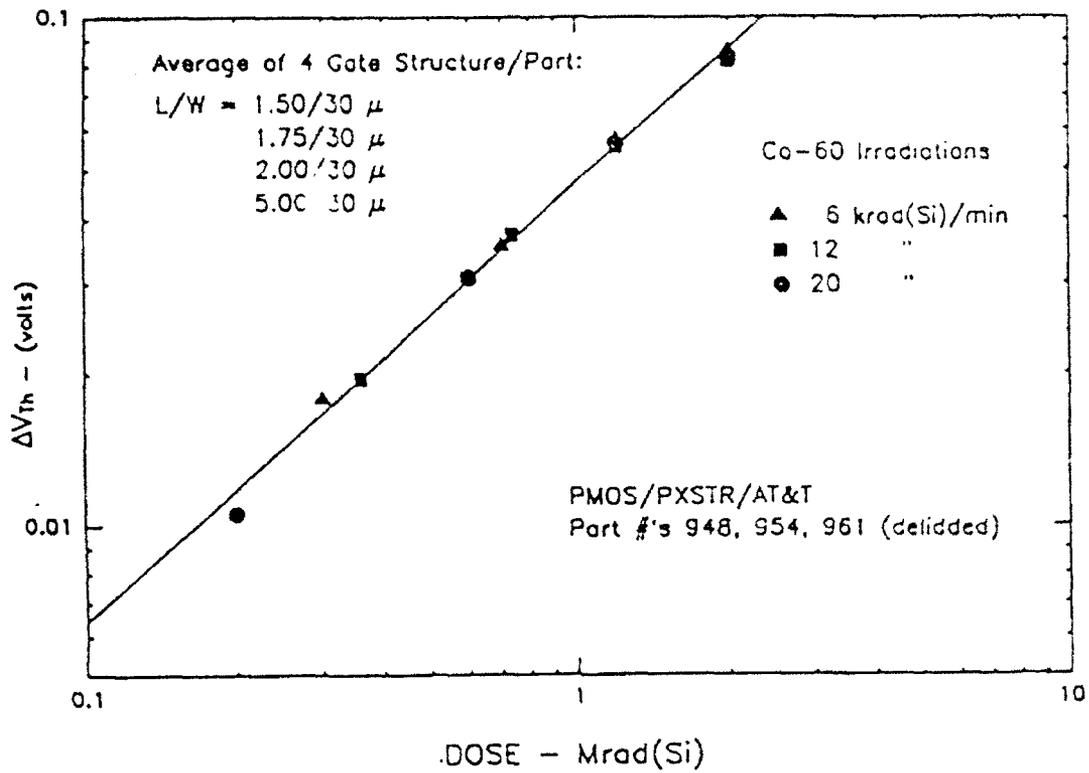


Figure 5: Co-60 Threshold Shift Data

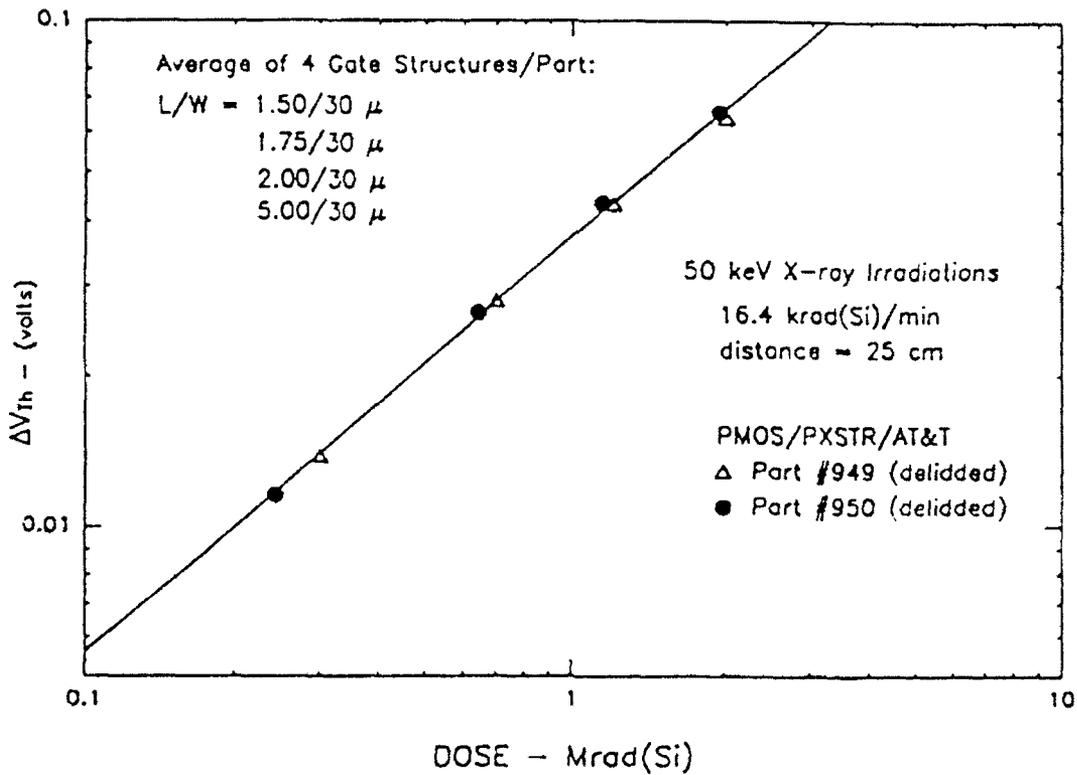


Figure 6: X-ray (10keV ave.) Threshold Shift Data

CONCLUSIONS

As shown in Table I, the calculated radiation intensity and diode current using the theoretical X-ray tube spectrum combined with spectral attenuation calculations due to air shielding and other filter materials was accurate to within 20% when compared to the response of a calibrated SEMIRAD detector. These results confirm that the analytical techniques used in this program lead to accuracies required by MIL-STD-883 (=20%) for radiation effects studies.

PHOTCOEF calculations of silicon dose deposition using the intensity/spectral model up to 50 keV also prove to meet MIL-STD-883 accuracies when compared to the experimental measurements using the calibrated Aracor PIN diode as shown in Table II.

Table II

Anode Voltage (KV)	Tube Current (mA)	Distance (cm)	Calculated Dose Rate (rad(Si)/sec)	Measured Dose Rate (rad(Si)/sec)	Error %
50	25	25	540	680	+20.5
30	25	25	325	344	+5.7
20	25	25	187	150.5	-2.1

This data adds confidence that the PIN diode in conjunction with the PHOTCOEF transport and deposition calculations will provide acceptable dosimetry calculations for a variety of materials and test geometries.

Figures 5 & 6 show little device-to-device variation (error bars fall inside the dots) in the sensitivity curves for the PMOS transistors exposed in Cobalt-60 or 50 keV X-rays. This preliminary data shows promise for application of these device structures as dosimeters.

Furthermore, they provide a direct correlation to Cobalt-60 as discussed in the referenced ASTM guidelines [1].

The confidence associated with the procedures developed in this effort will be extended into a complete calibration of the X-ray test chamber. At the present time, the dosimetry, accurate to within 10% and the analysis techniques, accurate to within 20%, can be applied with confidence while operating the X-ray source from 8-50 keV. Additional work is required to extend the complete experimental characterization of the facility over a broader range of X-ray energies. A later paper will report on ongoing work in this area.

References

- (1) L. Palkuti and J. LePage, "X-ray Wafer Probe For Total Dose Testing," IEEE Trans. on Nuclear Sci., NS-29 (1982).
- (2) D.B. Brown, "Std. Guide For The Use Of An X-Ray Tester (10keV Photons) in Ionizing Radiation Effects Testing of Microelectronic Devices," ASTM 11A49, July 1992
- (3) D.B. Brown, J.V. Gilfrich and M.C. Peckerar, "Measurement and Calculation of Absolute Intensities of X-Ray Spectra," Jour. Applied Physics, Vol.46, No.10, Oct. 1975
- (4) J.C. Garth, "An Algorithm For Calculating Dose Profiles in Multi-Layered Devices Using A Personal Computer," IEEE Trans. on Nuclear Sci., NS-33, No. 6 Dec. 1986
- (5) R. Bellem and R. Pelzl, "Analytical Tools to Support Dosimetry Calculations for a Low-Energy X-ray Test Cell," Phillips Lab Tech Report (to be published).
- (6) Private communication, D. Arion, SAIC.

THIS PAGE INTENTIONALLY LEFT BLANK

ULTRAWIDEBAND ANTENNAS WITH LOW DISPERSION

Albert W. Biggs
Professor
Department of Electrical and Computer Engineering

University of Alabama in Huntsville
Huntsville, AL 35899

Final Report for:
Summer Research Program
Phillips Laboratory
Kirtland Air Force Base NM 87117-6008

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

September 1992

ULTRAWIDEBAND ANTENNAS WITH LOW DISPERSION

Albert W. Biggs
Professor

Department of Electrical and Computer Engineering
University of Alabama in Huntsville

Abstract

Ultrawideband (UWB) radars or impulse radars are characterized by very wide bandwidths and commensurate fine range resolution. Applications include target identification in the regime of resonant frequencies of target aircraft, foliage penetration, and ground penetration for subsurface detection.

These radars often use "chirped" or pulse compression to increase radiated peak powers and subsequent range capabilities. Many UWB radars are designed without adequate knowledge or ability to design antennas which maintain adequate pulse waveform. While the input pulse to an antenna may be a damped sinusoid of 10-20 nanoseconds, the output is often a dispersively damped sinusoid of 10-20 microseconds. This paper studies UWB antennas which will reduce pulse distortion by decreasing frequency dispersion.

ULTRAWIDEBAND ANTENNAS WITH LOW DISPERSION

Albert W. Biggs

INTRODUCTION

Many antennas are highly resonant, operating over bandwidths of only a few percent. Such "tuned," narrow-bandwidth antennas may be entirely satisfactory or even desirable for single frequency or narrowband applications. In many applications described below, however, wider bandwidths are required.

Ultrawideband (UWB) radars or impulse radars are characterized by very wide bandwidths, larger than 25 percent of center frequency, and commensurate fine range resolution. UWB radars are useful for target identification [1] in the regime of resonant frequencies of the target aircraft.

They also have applications in short range side looking airborne radar (SLAR) for foliage penetration and point defense against low-observable sea skimmers [1] because of increased range resolution available with pulse compression in synthetic aperture radar (SAR). Potential ECM and weapons applications of UWB techniques include pulse compression to increase radiated peak powers and developing UWB antennas to reduce pulse distortion by decreasing frequency dispersion.

A survey [2] of UWB and frequency independent antennas from 1945 to 1962 describes discone antennas, Archimedes spirals, equiangular spirals, logarithmic periodic antennas, and many versions of log spiral and log periodic antennas. These antennas are brought to the present with a review paper by R. C. Hansen [3].

UWB AND FREQUENCY INDEPENDENT ANTENNAS

The preceding applications can be achieved with UWB and frequency independent antennas. A basic UWB monopole antenna fed with a coaxial (unbalanced) transmission line is the volcano smoke antenna [4]. By gradually tapering the inner and outer conductors of a coaxial transmission line, an UWB antenna with an appearance like that of a volcano crater and a puff of smoke is created. The ratio of outer to inner conductor diameter is relatively constant with the gradual smooth transition from coaxial line to a radiating structure, providing an almost constant input impedance over wide bandwidths. Since radiation occurs from narrower regions at shorter wavelengths, the radiation pattern tends to be relatively constant. These properties make the volcano smoke antenna a basic UWB monopole antenna. This antenna was built and tested at Harvard University Cruft Laboratory, and the impedance bandwidth was very broad as anticipated.

A basic UWB dipole antenna has two conductors, each resembling an Alpine-type horn used by Swiss mountaineers, with a two conductor (balanced) transmission line feed. It is called the twin Alpine horn antenna [4], which resembles a pair of Alpine type horns used by Swiss mountaineers. In the twin Alpine Horn antenna, a uniform transmission line section at the left gradually opens out until the separation is a wavelength or more with radiation from the curved region forming a beam to the right. The conductor spacing diameter ratio is constant, making the characteristic impedance constant over a wide bandwidth. Since radiation occurs from narrower regions at shorter wavelengths, the radiation

pattern tends to be relatively constant. These properties make the twin Alpine Horn a basic UWB antenna.

A compact version of the twin Alpine horn has a double ridge waveguide as the launcher on an exponentially flaring two conductor balanced transmission line. The design in Fig. 3 includes features used by J. L. Kerr [5] and by D. A. Baker and C. A. Van der Neut [6]. The exponential taper is of the form

$$Y = K(1) \exp [K(2) X]$$

where $K(1)$ and $K(2)$ are constants. The exact curvature is not critical provided it is gradual. The fields are bound sufficiently close to the ridges that the horn beyond the launcher may be omitted. The design shown is a compromise with the top and bottom of the horn present but solid sides replaced by a grid of conductors with a spacing of 0.10 wavelength at the lowest frequency. The grid reduces the pattern width in the H plane, increasing the low frequency gain. The cylindrical end sections [7]-[8] reduce the back radiation and VSWR. Absorbers on the top and bottom of the ridges or horn also reduce back radiation and VSWR.

TEM horns [9]-[12], with flare angle and plate widths chosen so only TEM modes are guided by maintaining constant impedance, and reduce dispersion [13]-[14] found with broadband or frequency independent antennas. TEM horns are made by taking out opposite sides of a pyramidal horn so that a two conductor line is formed by isosceles copper triangles. The effects of resistive loading of the above TEM horns were measured by M. Kanda [15]-[16], who assisted Sandia engineers in designing a similar TEM horn, and M. C. Bailey [17], who used either end loading or distributed loading.

A short TEM horn with a continuous tapered resistive loading was developed at the Electromagnetic Fields Division, NIST, Boulder, Colorado, for transmission of picosecond pulses with minimum distortion. It was found to be broadband and nondispersive with a low VSWR. The receiving transient response of the resistive loaded horn indicates the waveform of a 70 ps impulse was well preserved.

An antenna [18] developed at Phillips Laboratory, Kirtland AFB, combines a parabolic reflector fed by a conical TEM wave launcher. This launcher supports a step-like TEM wave on two or more conical conductors leading from some apex to the edge of the parabolic reflector. The antenna requires careful design of (1) switches with individual encapsulation, which serve as sources, (2) lenses near switches to provide higher dielectric strengths and to focus the outgoing wave, (3) sulfur hexafluoride region because of high power levels, (4) feed and launcher, (5) terminating impedances, and (6) parabolic antenna surface.

A ridged horn antenna was fabricated at Phillips Laboratory. The design begins with transitions from coaxial lines to single and double ridged waveguide sections in the construction of broadband coaxial-line-to-waveguide transitions [19]. Antipodal fin line horn antennas were also fabricated at Phillips Laboratory. With pulses less than 2 nanoseconds, anticipated dispersion was mild and not unexpected.

Since wide bandwidth antennas were desired, frequency independent antennas [20] were studied. Log periodic dipole arrays (LPDA) were made [21] and fed transient input pulses. With an

input pulse of duration T . dispersion broadens the radiated pulse directly as the difference between the low and high frequency periods of the LPDA and inversely as the tangent of the half angle of the LPDA. Amplitude and phase transfer functions were measured to use matched filter techniques. Use of phase conjugate functions [22], which were experimentally verified [23], produce pulse compression similar to that in synthetic aperture radars [24].

Transient analyses and experimental verifications were made at Phillips Laboratory by A. W. Biggs of a dipole, a pyramidal horn, and a LPDA [25]. His analytical models are described in papers similar to the those of K. Langenberg for LPDAs [26] and L. Marin [27] for dipole antennas. Another highly dispersive antenna is the spiral antenna, which is also an UWB antenna for steady state conditions. Pulse compression techniques for LPDAs is described in papers by A. W. Biggs and P. Ranon [28] and V. A. Yatskevich and L. L. Fedosenko [29], where reduction of dispersive effects was done with pulse forming [28] and redesign [29].

Discussions with M. Kanda at NIST would result in use of his TEM horn facility for antenna pulse measurements. A. Ondrejka, highly qualified in pulse measurements, may assist in measurements [30]-[32]. Interest in low-observable sea skimmers will result in an UWB antenna with microwave-millimeter frequencies in the pulses to discriminate against scatter from ocean waves [33], so that the configuration will resemble a fruit basket proposed by C. E. Baum at Phillips Laboratory. The foliage penetration radar [33] can be satisfied with a 20 to 50 Megahertz monopole-square loop array with resistive loading [9]. This antenna array will discriminate

against the higher frequency vegetation clutter [33] clutter, and identify the larger targets below the foliage. The monopole, quarter wave with resistive loading, will transmit, and the square loop will receive with a different polarization. The two antennas can be combined. High power microwave (HPM) antennas will be formed for square arrays and parabolically fed reflector antennas, centered around the pulse width transmitted (actually the inverse of the pulse width).

TRAVELLING WAVE ANTENNAS

A travelling wave antenna is one where fields and currents on the antenna can be represented by one or more travelling waves. Radiation occurs at discontinuities; when such discontinuities are closely spaced or are continuous then the antenna modes are leaky waves and radiation takes place continuously. In most cases radiation is unidirectional. Examples of travelling wave antennas are helical antennas, dielectric rod antennas, and long wire antennas. When the antennas are incorrectly terminated, reflected waves exist and the combination of oppositely travelling waves create standing waves. Antennas which support standing waves are travelling wave structures.

Linear antennas such as monopoles, which usually support standing waves, become travelling wave antennas when a resistance is inserted a quarter wavelength from the end of the antenna [34]. However this design doesn't create wideband characteristics for short pulse transmission because the value and location of the loading of the resistor are frequency dependent.

A unique use of a wire antenna [35] for detecting buried objects has the same horizontal wire acting as both transmitter and receiver, so that a pulse applied at the end of the wire (about 0.1 meter above the ground surface is detected by a current probe further along the same wire. A second, later pulse reflected from the buried object is also detected further along the wire.

A better structure is a V-shaped dipole operating in an end-fire mode is described by Iizuka [36], who developed a theory for single frequency operation. This antenna will support a TEM travelling wave of constant phase velocity, and is known as a TEM horn antenna.

A practical form of the TEM horn antenna is described by Wohlers [37], who developed an antenna whose characteristic impedance varied along the boresight axis of the antenna. Correct tapering of the impedance results in minimal internal reflections, while end loading and a balance feed structure reduce discrete reflections. It has a shape like a fish or teardrop, fed at the head of the fish or wide end of the water drop, and a conventional TEM horn described earlier.

SUMMARY

The antennas that can be used successfully for subsurface radar applications and target recognition are limited. This is because of the wide bandwidth required and large wavelength in relation to preferred antenna size. For Frequency Modulated CW subsurface radars, the choice is wider than for the short pulse

applications. The equiangular spiral antenna is useful where an extended impulse response is acceptable, but under conditions when the impulse response must be more limited the continuously loaded dipole or travelling wave antenna is preferred.

Effects of radiation and leakage are also factors in design of low dispersion UWB antennas. These are described by A. W. Biggs and E. A. Baca [38], in a paper to be presented at the National Radio Science Meeting January 5-8, 1993, in Boulder, Colorado. The abstract appears in Appendix A. A related study is the radiation fields from two wire transmission line antennas with varying lengths, by A. W. Biggs and E. A. Baca [38], in another paper to be presented at the National Radio Science Meeting. The abstract appears in Appendix A.

APPENDIX A

RADIATION LOSSES FROM TEM TRANSMISSION LINE ANTENNA FEEDS

Albert W. Biggs* and Ernest A. Baca

Phillips Laboratory, Kirtland AFB NM 87117

Abstract

TEM transmission line antenna feeds discussed here are grouped into open wire lines, shielded wire lines, strip lines, and conical lines. Although pairs of conductors and multiplicities of conductors are open wire lines, only two wire lines are considered because the application is feeding short dipoles, half wave and full wave antennas, loop antennas, helical antennas, biconical antennas, and TEM horns. Although open wire lines are simple and economical, they are unusable at high frequency because of excessive radiation loss. Shielded wire lines, such as coaxial lines, overcome the radiation problem.

Strip transmission lines have conductors in the form of flat parallel plates or strips, but open structures also suffer radiation problems at higher frequencies. As the strip width approaches zero, the two wire line configuration is approached. As the width approaches infinity, losses approaches zero, but this is not realistic.

Several antenna feeds are associated with conical lines. One line has two conductors each resembling an Alpine-type horn used by Swiss mountaineers. When the curved conductors are straightened into regular cones, they become a V-shaped conical feed.

When aligned colinearly, a biconical feed is formed. If the lower cone angle increases to 180 degrees, a feed with a conical section above a ground plane is reached.

Other variations include a right cylindrical cone cut by two planes passing through the apex, forming a rectangular at the base of the cone. The two conductors can be the isosceles triangles in the two planes cut by the cone or the curved cone sections cut by the planes. These feeds provide wide band coupling to TEM horn antennas because a constant characteristic impedance is maintained as the feed cross section increases or decreases.

This evaluation was made because losses due to radiation from antenna feeds is usually neglected, but with higher frequencies in pulsed waveforms and high power microwave sources, these are no longer negligible. Losses are calculated for two wire, parallel strip, and conical TEM feed lines, with associated fields.

*Visiting Faculty at Phillips Lab until 15 Sept 92, then
Electrical and Computer Engineering Department, University of
Alabama in Huntsville, AL 35899.

APPENDIX B

RADIATION FIELDS FROM TWO WIRE TRANSMISSION LINE ANTENNAS

Albert W Biggs* and Ernest A. Baca

Phillips Laboratory, Kirtland AFB, NM 87117

Abstract

The family of antennas analyzed and described here originate with two parallel linear perfectly conducting, infinitely thin wires. Spacing between antenna elements is much less than the source wavelength. Excitation currents are constant along the wires, flowing in opposite directions.

Families of these antennas are formed by varying the length of the two wire lines. When length is small compared with the source wavelength, the antenna becomes a pair of infinitesimal or Hertzian dipoles. With longer lengths, they are similar to wave antennas (H.H. Beverage et al, AIEE Trans., 42, 215-266, 19 3, transmission line antennas, and horizontal antennas buried in Antarctic ice and snow (A.W. Biggs, IEEE Trans. AP-16, 201-208, 1968. Wave antennas, or Beverage antennas, are single long low wires above the earth's surface, first erected on Long Island and in Scotland in 1923.

Transmission line antennas are single wires above perfectly conducting ground planes so that the lines and their images form a two wire transmission line. These are encountered with coaxial feeds on aircraft, where outer conductors terminate on the aircraft surface and the inner conductor continues slightly and then

bends to be parallel with the surface. Terminations may be short or open circuits and mismatched or matched loads. Matched loads simulate constant current excitation, while other loads introduce standing waves.

When lengths increase to several wavelengths, broadside or main lobes increase in magnitude, side lobes increase in number and magnitude, and mainlobe and sidelobe beamwidths decrease. As antenna lengths reach infinity, mainlobe and sidelobe beamwidths reach zero, and the mainlobe increases in the form of Dirac Delta functions.

Antenna patterns are presented for the above varying lengths.

*Visiting Faculty at Phillips Lab until 15 Sept 92, then Elect and Compu Engr Department, University of Alabama in Huntsville, Huntsville, AL 35899

REFERENCES

- [1] OSD/DARPA UWB Radar Review Panel, 'Assessment of Ultra-Wide (UWB) Technology,' Battelle Tactical Technology Center. Columbus, Ohio, July 13, 1990.
- [2] J. D. Dyson, 'A Survey of the Very Wide Band and Frequency Independent Antennas,' J. Research NBS, Part D, vol. 66D, pp. 1-6, Jan.-Feb. 1962.
- [3] R. C. Hansen, 'HPM Antennas: Challenges and Wideband Capabilities,' Proceedings of the 5th National Conference on HPM Technology, West Point Military Academy, NY, June 1990.
- [4] J. D. Kraus, Antennas, 2nd ed. New York: McGraw-Hill, 1988, pp. 692-710.
- [5] J. L. Kerr, 'Short axial length broad-band horns array,' IEEE Trans. Antennas and Propagat., vol. AP-21, pp. 710-714, Sept. 1973.
- [6] D. E. Baker and C. van der Neut, 'A compact, broadband, balanced transmission line antenna derived from double ridged waveguide,' Proc. of IEEE APS Symposium. Albuquerque, NM, pp. 568-570, 1982.
- [7] C. W. Chuang and W. D. Burnside, 'A diffraction coefficient for a cylindrically truncated planar surface,' IEEE Trans. Antennas and Propagat., vol. AP-28, pp. 177-182, Mar. 1980.
- [8] W. D. Burnside and C. W. Chuang, 'An aperture matched horn design,' IEEE Trans. Antennas and Propagat., vol. AP-30, pp. 790-796, July 1982.

- [9] M. Kanda, 'Transients in a resistively loaded linear antenna compared with those in a conical antenna and a TEM horn,' IEEE Trans. Antennas and Propagat., vol. AP-28, pp. 132-136, Jan. 1980.
- [10] E. A. Theodorou, M.H. Gorman, P. H. Rigg, and F. N. Kong, 'Broadband pulse-optimized antenna,' IEE Proc., Part H. vol. 128, pp. 124-130, June. 1981.
- [11] S. Evans and F. N. Kong, 'TEM horn antenna: Input reflection characteristics in transmission,' IEE Proc., Part H. vol. 130, pp. 403-409, Oct. 1983.
- [12] C. L. Bennett and G. F. Ross, 'Time domain electromagnetics and its applications,' Proc. IEEE, vol. 66, pp. 299-318, Mar. 1978.
- [13] J. K. Pulfer, 'Dispersive properties of broad-band antennas,' Proc. IRE, vol. 49, p. 644, Mar. 1961.
- [14] C. M. Knop, 'On transient radiation from a log periodic dipole array,' IEEE Trans. Antennas and Propagat., vol. AP-18, pp. 807-808, Nov. 1970.
- [15] M. Kanda, 'A relatively short cylindrical broadband antenna with tapered resistive loading for picosecond pulse measurements,' IEEE Trans. Antennas and Propagat., vol. AP-26, pp. 439-447, May 1978.
- [16] M. C. Bailey, 'Broad-band half-wave dipole,' IEEE Trans. Antennas and Propagat., vol. AP-32, pp. 410-412, May 1972.
- [17] M. Kanda, 'The effects of resistively loading of 'TEM' horns,' IEEE Trans. EMC, vol. EMC-24, pp. 245-255, May 1982.

- [18] C. E. Baum, 'Radiation of impulse like transient fields,'
Air Force Weapons Laboratory, Kirtland Air Force Base, NM,
Sensor and Simulation Note 321, Nov. 1989.
- [19] G. L. Ragan, Microwave Transmission Circuits, MIT Radiation
Series Volume 9, Boston Technical Publishers, pp. 355-364,
1964.
- [20] V. H. Rumsey, Frequency Independent Antennas. New York:
Academic Press, 1966.
- [21] D. E. Isbell, 'Log Periodic Dipole Arrays,' IRE Trans.
Antennas and Propagat., vol. AP-8, pp. 260-267, May 1960.
- [22] P. Van Etten, 'EM field compression by dispersive broadband
antenna,' IEEE/APS Symposium, pp. 350-353, June 1974.
- [23] P. J. Steibach and P. Van Etten, 'An Experimental Demons-
tration of Pulse Compression Antennas', IEEE/APS Symposium,
Seattle, WA, pp. 216-219, June 18-22, 1979.
- [24] A. W. Biggs and J. M. Jordan, 'Simulation of chirp pulses and
effect of doppler shifts on backscattered pulses,' Proc. of
IGARRS 89 Symposium, Vancouver, Canada, pp. 1720-1723,
June 10-14, 1989.
- [25] A. W. Biggs, 'The Transient Response of a Pyramidal Horn
Antenna,' International IEEE-APS Symposium, London, Ontario,
Canada, June 24-28, 1991.
- [26] K. J. Langenberg, 'Transient Fields of Linear Antenna Arrays,'
Applied Physics, vol. 20, pp 101-118.
- [27] L. Marin and T. K. Liu, 'A Simple Way of Solving Transient
Thin-Wire Problems,' Radio Science, vol. 11, pp. 149-155,
Feb. 1976.

- [28] A. W. Biggs and P. Ranon, 'Pulse Compression on Log Periodic Dipole Array Antennas,' Proceedings of the 5th National Conference on HPM Technology, West Point Military Academy, NY, June 1990.
- [29] V. A. Yatskevich and L. L. Fedosenko, 'Antenna Dlya Izlucheniya Sverkhshirokopolosnykh Signalov,' Radioelektronika, vol. 29, no. 2, pp. 69-74, 1986.
- [30] A. R. Ondrejka and P. A. Hudson, 'Measurement Standards for Low and Medium Peak Pulse Voltages,' J. Research NBS, Part C, vol. 70C, pp. 13-18, Jan.-Mar. 1966.
- [31] A. R. Ondrejka, 'Peak Pulse Voltage Measurement,' Proc. IEEE, pp. 882-885, June 1967.
- [32] A. R. Ondrejka and others, 'Applications of Time-Domain Methods to Microwave Measurements,' IEE Proc., Part H, vol. 127, pp. 99-106, April 1980.
- [33] A. W. Biggs, 'Digital Simulation of Terrain Backscattering Cross Sections,' Proc. of IGARRS 89 Symposium, Vancouver, Canada, pp. 1000-1003, June 10-14, 1989.
- [34] E. E. Altshuler, 'The travelling-wave linear antenna,' IRE Trans. Antennas and Propagat., vol. AP-9, pp. 324-329, July 1961.
- [35] P. Degauque and J. P. Thery, 'Electromagnetic subsurface radar using the transient field radiated from a wire antenna,' IEEE Trans. Geoscience Electronics, vol. GE-24, pp. 805-812, May 1986.

- [36] K. Iizuka, "The travelling-wave V-antenna and related antennas," IEEE Trans. Antennas and Propagat., vol. AP-15, pp. 236-243, July 1967.
- [37] R. J. Wohlers, "The GWIA, an Extremely Wide Bandwidth Low Dispersion Antenna," Abstracts of the 20th Symposium on USAF Antennas R & D Progress, Oct. 1970.
- [38] A. W. Biggs and E. A. Baca, "Radiation Losses from TEM Transmission Line Feeds," NRC/URSI Joint Meeting, University of Colorado, Boulder, Jan. 5-8, 1993.
- [39] A. W. Biggs and E. A. Baca, "Radiation Fields from Two Wire Transmission Line Antennas," NRC/URSI Joint Meeting, University of Colorado, Boulder, Jan. 5-8, 1993.

OPTICAL ANGLE-ANGLE DOPPLER IMAGING

J.K. Boger
Assistant Professor
Laser Optical Engineering Technology

Oregon Institute of Technology
Klamath Falls, Oregon 97601

Final Report for:
AFSOR Summer Research Program
Phillips Laboratory
LIMI

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

September 1992

OPTICAL ANGLE-ANGLE DOPPLER IMAGING

J.K. Boger
Assistant Professor
Laser Optical Engineering Technology
Oregon Institute of Technology

ABSTRACT

A coherent imaging experiment has been done where the modulus and phase were measured directly using a heterodyne technique. Information gathered in the transform plane was manipulated to improve the final image over an image constructed from a distorted wave front. Differences between up-link and down-link distortions were investigated and experimental results obtained are presented. Doppler information was also obtained and analyzed for translating and rotating targets. Results of averaged speckle images obtained from Doppler broadened data are presented.

HETERODYNE CONCEPTS

Coherent imaging is a process of collecting target data in the Fourier plane and then spatially transforming it to construct a final image. Normally, a positive lens does this spatial transform, resulting in a real image. But the coherent imaging process gives us the privilege of working with the data for our benefit. One such benefit explored here is the attempt to remove phase distortions from the wave front for a clearer image. Also of interest is the fact that the Fourier plane contains additional target information not recovered by a lens. Information such as target velocity is lost when imaging with a conventional glass lens, but may be recovered using the proper techniques. These techniques include recovering the modulus and phase of the wave front in frequency space. There are several methods of achieving this goal.² Here the field was measured using a heterodyning process of data collection.

One of the great advantages of heterodyne is the direct measurement of phase. It is a clever method used to measure the phase of a wave whose frequency is far beyond the reach of direct measurement. When dealing with the ultra-high frequencies of light, we must resign ourselves to the use of intensity detectors which implies the loss of phase information. This is due to the fact that intensity (or more correctly, irradiance) is proportional to the square of the electric field, U_T .

$$I = \eta U_T^*(\vec{r}, t) \cdot U_T(\vec{r}, t) \quad 1$$

where

$$U_T^*(\vec{r}, t) = E(\vec{r}) e^{-i(\omega t + \varphi_T)} \quad 2$$

Here the frequency information is represented by ω , φ carries the phase and η is a proportionality constant. The position vector r locates a point in the pupil plane. Evidently the electric field, U_T had the phase information in it but the multiplication of the phase with its complex conjugate effectively loses this information.

Heterodyne circumvents the loss of phase by mixing two waves at slightly shifted frequencies and recovering the cross product produced after squaring the superposition. This is better explained with the relatively simple mathematics of superposition. Equation 3 is the superposition of the two fields.

$$U_{TOTAL}(\vec{r}, t) = U_T(\vec{r}, t) + U_{LO}(\vec{r}, t) \quad 3$$

where

$$U_{LO}(\vec{r}, t) = E(\vec{r}) e^{i(\omega_{LO}t + \varphi_{LO})} \quad 4$$

Squaring the superposition given in equation 3 gives

$$I_{TOTAL} = I_T + I_{LO} + \eta U_T^* U_{LO} + \eta U_{LO}^* U_T \quad 5$$

The first two terms in the intensity are of little interest but the last two terms represent the heterodyne signal. This heterodyne signal can be written explicitly in terms of modulus and phase as in equation 6.

$$S_H(\vec{r}, t) = \eta E_{LO}(\vec{r}) E_T(\vec{r}) \{ e^{i(\omega_{IF}t + \varphi)} + e^{-i(\omega_{IF}t + \varphi)} \} \quad 6$$

Here the intermediate frequency ω_{IF} , is the difference between the local oscillator frequency ω_{LO} and the target frequency ω_T . Similarly, the net phase φ is the difference in phase between the two mixing fields. The physical effect is to produce an oscillating intensity at the detector. The heterodyne signal can be Fourier transformed in time to expose the modulus and phase of the mixed field in the detector plane.

$$S_H(\vec{r}, +\omega_{IF}) = \eta E_{LO}(\vec{r}) E_T(\vec{r}) \{ e^{i(\varphi)} \} \quad 7$$

$$S_H(\vec{r}, -\omega_{IF}) = \eta E_{LO}(\vec{r}) E_T(\vec{r}) \{ e^{-i(\varphi)} \} \quad 8$$

The heterodyne process is repeated pixel by pixel over the entire 64x64 detector array to generate two dimensional maps of the modulus and phase. The maps are then multiplied together to generate a field which is inverse spatially transformed to construct an image. The net result of this process is shown in Figure 1, precisely as it is displayed on the Trapix monitor. The single speckled image was produced by a reflective metallic sharks tooth which is described in the next section.

After the modulus and phase information has been determined, data analysis can begin. Here data analysis refers to the manipulation and interpretation of the phase in frequency space. A simple manipulation could be to take the complex conjugate of the phase to reverse the final image. In this way it is possible to flip the sharks tooth image to match the incoherent image in Figure 4. An interpretation of the data would be measuring the linear

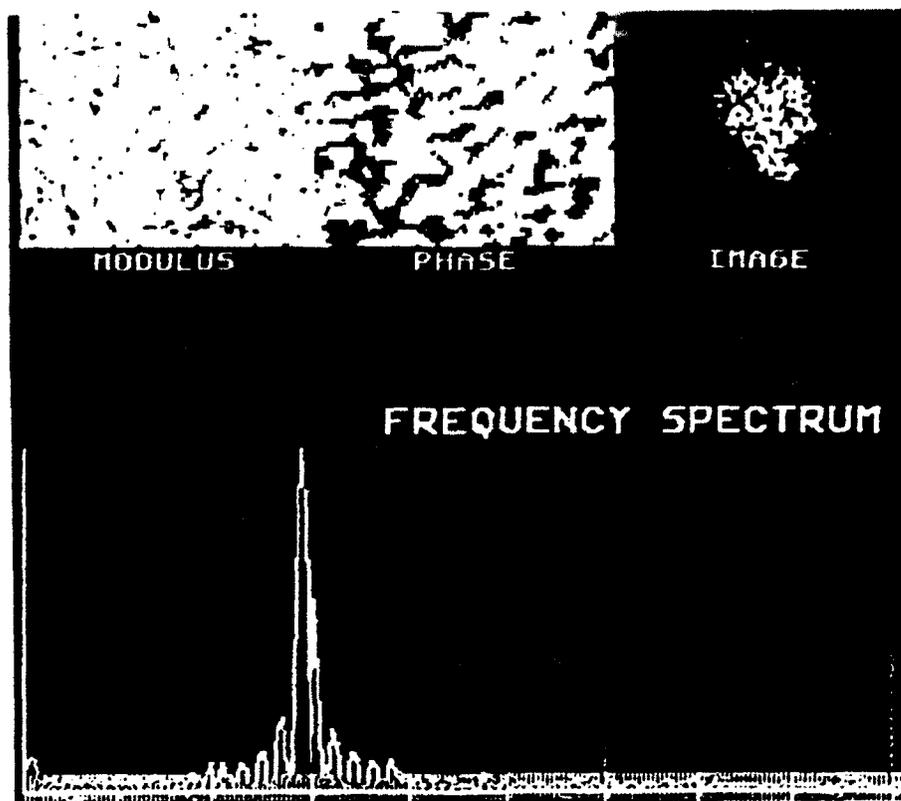


Figure 1. Speckled imaging results using the heterodyne process.

phase shifts to determine a position shift in the calculated image. This concept is an important foundation in this work and merits further explanation.

The key to successful data manipulation is to remember how the changes will be spatially transformed. The Fourier transform of a linear phase is a spatial position shift in the result.

$$F^{-1}[e^{-i(2\pi x_0 \Omega)}] = \delta(x - x_0) \quad 9$$

It is important to keep in mind that Ω is a spatial frequency and not temporal. The linear phase is revealed in the data as waves of tilt. Figure 2 illustrates the difference between a small and large linear phase. The data for this figure was taken from Doppler broadened data for clarity. The images are only portions of a reflective target which are clearly spatially shifted relative to each other. With the linear phase shift evidenced in the Figure, it is only natural to ask what are causes for this effect.

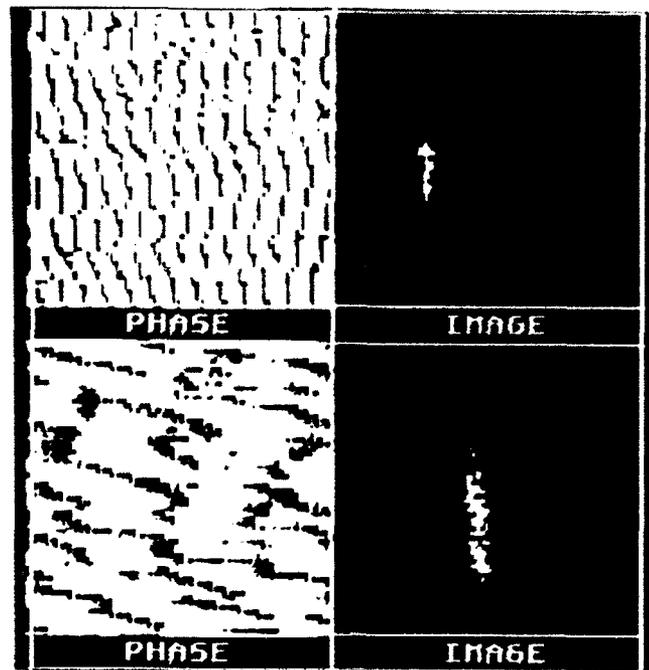


Figure 2. Large spatial shift relative to a smaller shift

There are three simple ways to produce a linear phase. The first and simplest way is to reposition the target in the illuminating beam. The second method is to introduce a misalignment between the target beam and the local oscillator. As an experimental technique it is normal to adjust the position of the target within the illuminating beam and re-center the image by adjusting the local oscillator alignment. Finally, it is possible to force the spatial shift by multiplying the phase directly by a linear spatial frequency. The important point to keep in mind is that the experimenter is in control of the spatial position of the image through the phase manipulations.

EXPERIMENTAL ARRANGEMENT

The experimental arrangement used was similar to the heterodyne arrangement used one year ago. Figure 3 diagrams the basic optical arrangement used for both the transmissive and reflective targets. Many of the experimental details were reported on last year and iteration here will be minimized.³

The laser used was an argon-ion laser of 500 mW power and operating at 514 nm single line. Two acousto-optic modulators were responsible for separating the frequency of the two beams by 8 hertz. The CCD camera was reset at a rate of 64 hertz and each frame consisted of 64x64 pixels, 15 μ m on center. The maximum reset rate for the camera was 128 hertz. Data acquisition was achieved using the Trapix image processing environment. The Trapix

had a maximum of 512 frames of data storage capability. The transform lens used was a 50 cm focal length lens with 63 mm diameter.

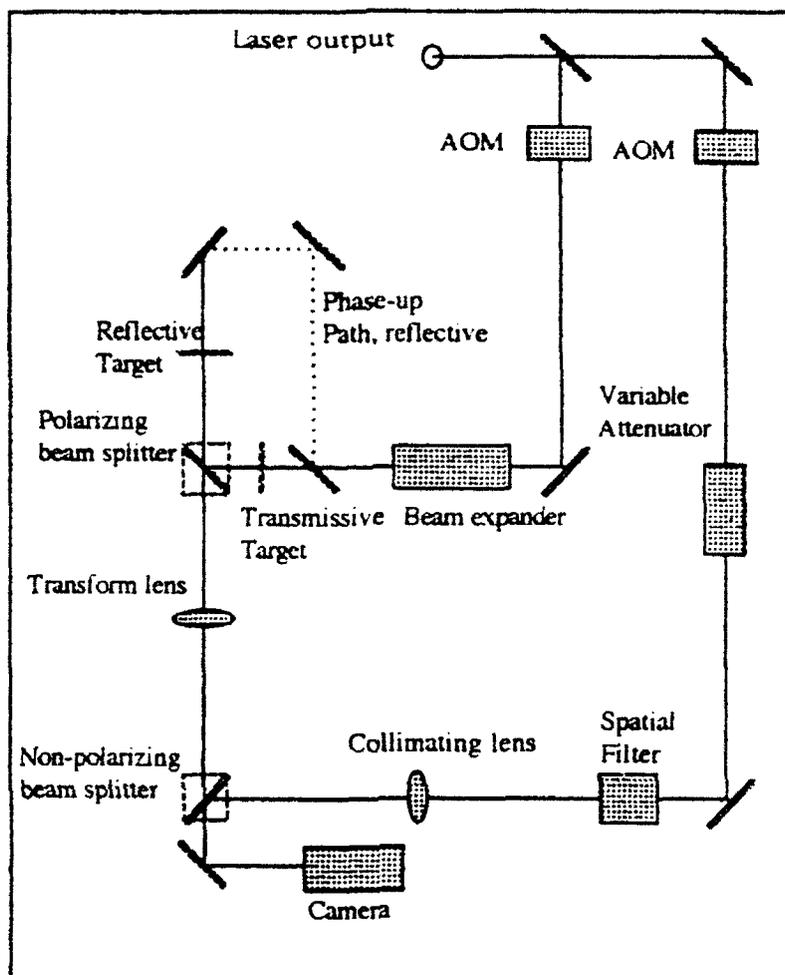


Figure 3. Experimental arrangement.

The speed of the lens is slower as mounted in the experiment due to a glare stops acting as aperture stops. The glare stop was placed near the lens and had a diameter of 50 mm.

Two optical changes were made in the arrangement as compared with the previous heterodyne experiment. First, the large beam splitter used to recombine the two arms of the interferometer was replaced with a non-polarizing beam splitting cube. This was done to reduce the interference fringes arising from multiple reflections in the flat beam splitter. The second change concerns the placement of the transform lens. Last year the lens was placed after the recombination of the optical fields. This year the was placed before the recombination. This necessitated the use of the collimation lens after the spatial filter.

The arrangement worked well for the experiment with few exceptions. One notable problem was a very low frequency noise in the local oscillator. The source of this fluctuation was not the laser. This suggests that the problem is in the acousto-optic modulator. However, if the modulator is the problem, it was not due to the frequency input which was very stable. Instead, the diffraction efficiency of the crystal seemed to vary as if the incident angle was changing due to thermal variations. The crystals were liquid cooled.

Finally the targets used should be described. Both the transmissive and reflective targets used were the same as those used last year. The primary transmissive target was the NRC RES-1 resolution bar target. Speckle was achieved by covering the 0.5 mm bar target with plastic tape. The incoherent image of the transmissive target, as well as a typical speckled image and average speckle image are shown in Figure 4a. Figure 4b shows the same sequence of images for the reflective target used which was the 6 mm wide metallic sharks tooth.

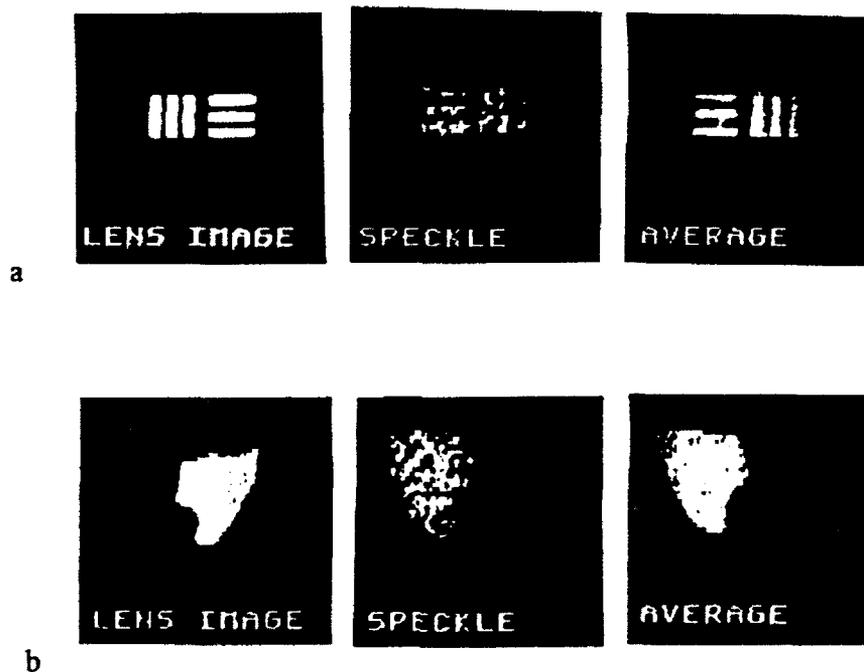


Figure 4. a) Transmissive bar target. b) Reflective sharks tooth target.

PHASE-UP

The name given to the process of removing phase distortions is phase-up. This work has been done by other investigators in the field and is repeated here to complete the imaging procedures.⁴ A simple outline of the mathematics behind phase-up greatly facilitates the

discussion. A more detailed treatment has been presented by Dr. R. Anderson in a report submitted to the Phillips lab Imaging branch (LIMI). The discussion to follow considers only static targets, either transmissive or reflective.

Our investigation of phase distortions begins by mathematically quantifying them in the electric field expressions.

$$U_T(\vec{r}, t) = E_T(\vec{r}) e^{i(\varphi' + \gamma + \omega_T t)} \quad 10$$

$$U_{LO}(\vec{r}, t) = E_{LO}(\vec{r}) e^{i(\varphi'' + \omega_{LO} t)} \quad 11$$

The difference between these expressions and the electric fields given in the last section is that the characteristics of the target are represented by $e^{i\gamma}$ while the distortions in the target and local oscillator wave fronts are represented by $e^{i\varphi'}$ and $e^{i\varphi''}$ respectively. Distortions in the local oscillator are an experimental parameter and can therefore be controlled. Distortions in the target wave front are usually an unavoidable consequence of imaging through the atmosphere. Consequently the principal effort in phase-up was to remove target distortions.

The heterodyne process combines the phase from each wave front giving a net phase distortion seen in the heterodyne signal.

$$S_H(\vec{r}, t) = \eta E_{LO}(\vec{r}) E_T(\vec{r}) \{ e^{i(\omega_{IF} t + \varphi + \gamma)} + e^{-i(\omega_{IF} t + \varphi + \gamma)} \} \quad 12$$

Here, as in equation 6, $\varphi = \varphi'' - \varphi'$ and ω_{IF} is the intermediate frequency. The information we seek is in $e^{i\gamma}$. When the temporal Fourier transform is done we are left with a modulus and phase.

$$S_H(\vec{r}, +\omega_{IF}) = \eta E_{LO}(\vec{r}) E_T(\vec{r}) e^{i\gamma} e^{i\varphi} \quad 13$$

The phase distortion term in equation 13 is in bold to emphasize that we have isolated the negative effect on the wave front and consequently should be able to remove it. The idea behind eliminating $e^{i\varphi}$ is identical to the way phase information was lost when detecting light with an intensity detector. That is, multiply the phase by the complex conjugate of the same term. Our goal at this point is to measure $e^{i\gamma}$ without the target information, $e^{i\varphi}$. This was accomplished by replacing the target with a transmissive pinhole.

The size of the pinhole was calculated so that the transform of it was a very broad sombrero function in the Fourier plane. The pinhole simply replaced the target in the transmissive imaging experiment. For the reflective experiment, it was necessary to redirect the target beam as shown by the dotted line in Figure 3. An alternative method of phasing-up was tried by replacing the target with a small ball bearing. This method produced very poor phase calibration due to spherical aberrations and micro structure of the ball. The temporal transform of the pinhole data is given by equation 14.

$$S_{HP}(\vec{r}, -\omega_{IF}) = \eta E_{LO}(\vec{r}) E_P(\vec{r}) e^{-i\phi} \quad 14$$

The computer algorithm which is responsible for calculating the field does so by discretely calculating the modulus and phase. Consequently it was a simple matter to isolate $e^{i\phi}$. Notice that the negative frequency component of the transform was chosen since it represents the complex conjugate of the phase distortions sought. Simple multiplication of this measured distortion with equation 13 removes the unwanted phase and leaves the target information in unaltered. An additional effect may be that the pinhole phase has a linear phase shift associated with it in addition to possible phase distortions. Figure 5 shows the transmissive bar target image constructed after collecting pinhole phase information. No phase distorter was used. The pinhole was intentionally shifted to produce a linear phase, thereby off centering the final image. This is the effect of multiplying the field by a linear phase as discussed in the heterodyne concepts section.

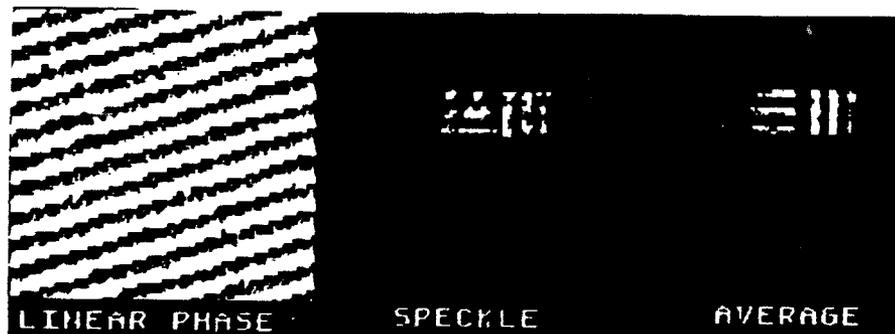


Figure 5. Off centering the final image with a linear phase-up.

Removing atmospheric distortions are the principal interest in the area of advanced coherent imaging. Distortion in the wave front can occur when the illuminating beam propagates from the source to the target. This is referred to as up-link distortions. So called down-link distortions occur when the coherent radiation propagates from the target to the

detector. Each distortion could be investigated separately and was. It was only a matter of placement of the phase distorter in the experiment.

The first result shown focuses on the up-link distortions. The distorter was a thick piece of a plastic sandwich bag placed in the target beam before the reflective target. After extensive investigation, it was determined that the up-link distortion had very little effect on the averaged image. A simple explanation for this is that the distorter does little more than introduce an arbitrary phase in the wave front exactly as the microscopic structure of the target does. This is not a negative effect since there is no target information in the wave front to disturb. We could not hope to separate the phase variations caused by microstructure versus up-link distortion. Luckily, as the experiment suggests, it makes no difference!

Down-link distortions are a completely different matter. Keeping in mind that the wave front is carrying target information on the down-link, any distortions will obscure the image sought. Figure 6 displays the results from the down-link experiment. The conditions and procedures were the same as in the up-link experiment. Clearly the top image, which is the uncorrected image has been obscured by the distorter while the corrected bottom image is much better.

DOPPLER IMAGING

Possibly one of the greatest benefits associated with the heterodyne process is the ability to retrieve and process Doppler broadened data. Figure 7 is the frequency plot for a static target. In this plot the peak is very narrow and centered on 8 hertz, the frequency separation between the A-O modulators. Figure 8 is the frequency spectrum of a rotating and translating target. Clearly the Doppler effect is present. This section of the report discusses the basic concept of the Doppler effect and presents the significant results obtained. Doppler shifts in the target frequency occur when the illuminating beam reflects off of a moving target. If the movement is purely radial, the effect should be a simple shift in the narrow frequency curve away from the 8 hertz position. The shift in frequency is evident in the Doppler equation which reduces to

$$v = \Delta f \lambda$$

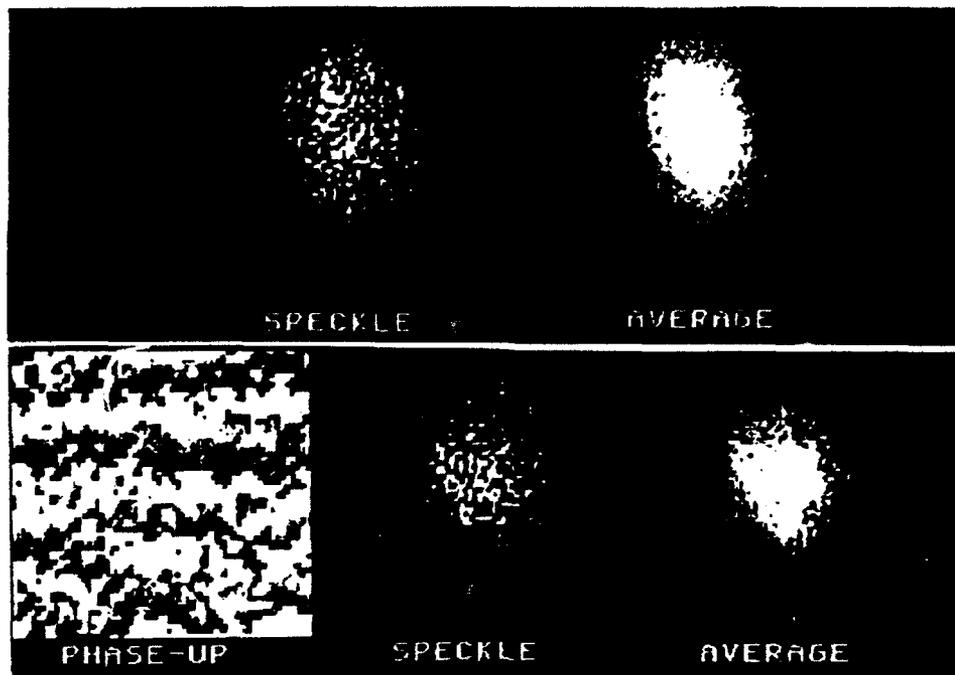


Figure 6. Results of down-link phase-up corrections in the reflective target.

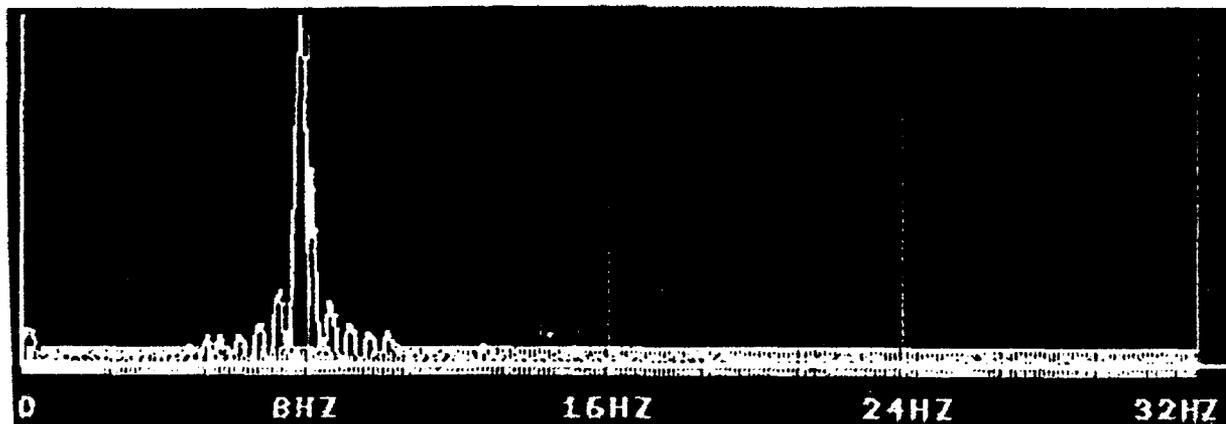


Figure 7. Frequency spectrum from a static target.

Recalling that the illuminating laser was operating at a wavelength of 0.514 mm, it is clear that very small velocities will create very large frequency shifts. It is because of equation 10 that the velocities of the laboratory targets were always on the order of microns per second. It is a fair question to ask what this Doppler shift looks like mathematically.

The frequency of the beam suffers a shift upon reflection from the target. This frequency shift is part of the phase that mixes with the local oscillator wave front yielding a shifted heterodyne signal given as

$$S_{HD}(\vec{r}, t) = \eta U_{LO}(\vec{r}) U_T(\vec{r}) \{ e^{i(\gamma + \omega_{IF}t + \omega_{DT})} + e^{-i(\gamma + \omega_{IF}t + \omega_{DT})} \} \quad 16$$

If the motion of the target were restricted to uniform motion along the line of sight, the frequency curve would look exactly as it does in Figure 7 except that the peak would be shifted away from 8 hertz. It should also be noted that the velocity of the target is not restricted to small values since the shape of the curve remains narrow. To compensate for large in-line motions it is possible to separate the intermediate frequency arising between the modulators. This was done in the lab using a Klinger motorized stage to move the target at 5 mm/sec while the modulators were separated by 10,000 hertz. The frequency spectrum was recovered as well as a processed image. Unfortunately few targets have such restricted motions.

The more challenging target motion is rotational. In this motion, the frequency spectrum broadened as seen in Figure 8.

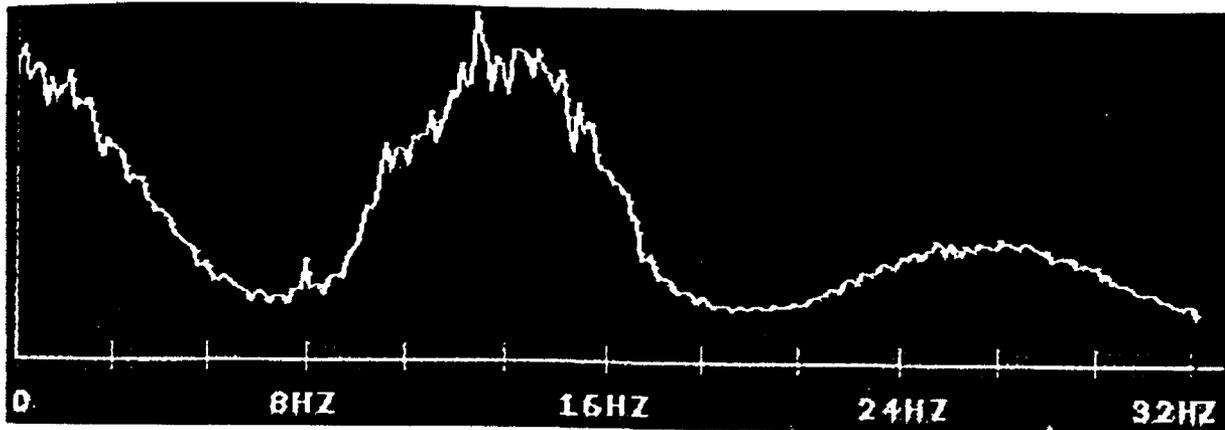


Figure 8. Frequency spectrum from a rotating and translating target.

Mathematically, the signal has now been distributed through a range of frequencies. The temporal Fourier transform yields a term which is genuinely a function of frequency unlike the static case where all of the information was at a single frequency.

$$S_{HD}(\vec{r}, \omega) = \eta U_{LO}(\vec{r}) U_T(\vec{r}) \{ e^{i(\gamma + \omega t)} \} \quad 17$$

Notice that this transform differs from the transform presented in the Heterodyne Concepts section in two important respects. First, the independent ω variable no longer carries an IF subscript since it is representing a range of frequencies. Second, the phase of the target, $e^{i\gamma}$ is subscripted by ω to emphasize the fact that the phase changes as the Doppler shifted

frequency changes. This range of frequencies is a direct result of having different regions of the target moving with different velocities. In the case of a purely rotating target, the center of the target has no net motion, consequently no Doppler shift. But the edges of this target would be at a maximum velocity relative to any other point in the target and this region would be responsible for the maximum Doppler shift. In other words, if the target were divided into several discrete sections, each would carry its own unique modulus and phase.

The range of frequencies the data can be processed at restricts the region of the target which can be imaged. It is possible to develop a formula which predicts the physical size of a region on the target which has the proper velocity. Figure 9 shows the geometry for this situation and equation 18 relates δr to δf .

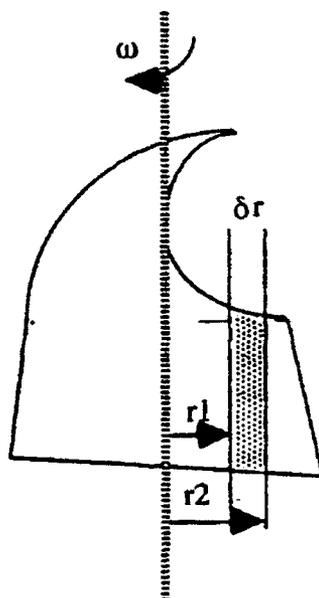


Figure 9. Target region imaged by discrete frequency

$$\delta r = \frac{\delta f \lambda}{\omega}$$

18

Here δf represents the smallest frequency division we have processed at. Usually this is 1/8th hertz which corresponds to 512 frames of data. This concept of relating the number of frames of data to the width of a strip in a target which is important to image reconstruction and warrants some explanation.

To begin with, it is necessary to know the frequency range where our data may lie. Since the camera is acquiring data at a rate of 64 hertz, the smallest time interval it can separate is about 0.0156 seconds. When this time interval is transformed into frequency

It is clear that our frequency range is 64 hertz. This means that the actual frequency lies between -32 hertz and +32 hertz. To find the smallest interval in frequency space, it is necessary to know the number of data points which have been processed. This leads to the concept of the frequency bin. Frequency bin size is inversely related to the number of data points processed. In this experiment the maximum number of frames was 512. If all 512 frames of data were processed, then the frequency spectrum would have 8 calculated points per hertz. It then follows that 32 frames would allow for one point per 2 hertz in the spectrum. Consequently, 512 frames yields a very narrow strip of the image when one frequency component is processed while 32 frames images a larger portion of the target. This is clearly evident in Figure 10 where four strips of an image are produced from the same frequency value but a different number of frames processed. The left image was produced after 32 frames were processed while the right image was produced after processing 512 frames.

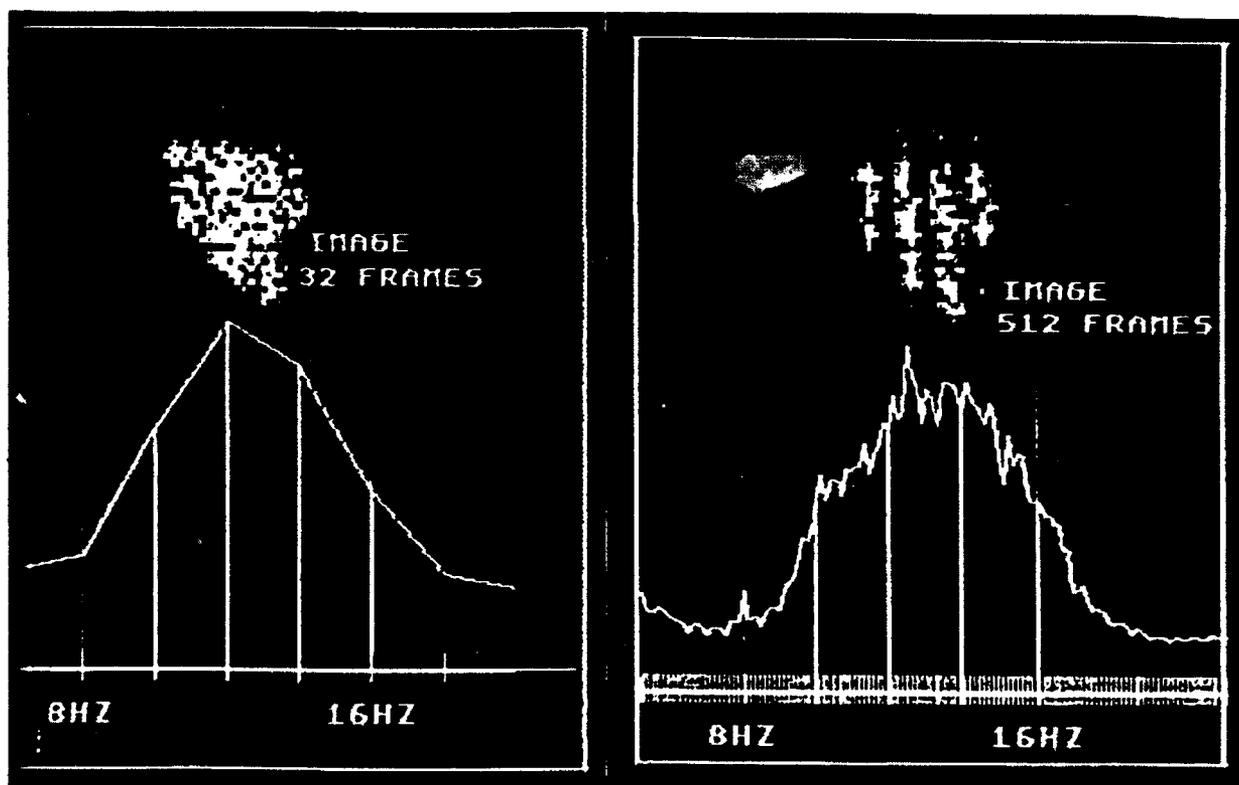


Figure 10. Image strips from a) 32 frames, b) 512 frames

As a consequence of the frequency bin size, the range of the target which can be imaged at one frequency bin is restricted. This means that different positions in the target are

imaged at different frequencies. A shift in the position of an image is produced by a linear phase bias. If the linear phase is large, the imaged portion of the target is considerably off center. If the linear phase is small, the imaged portion of the target is near the center. This linear phase shift is easily seen in the phase map as a number of waves of tilt in the wave front. Figure 11 is a picture of the last three calculated modulus, phase, image and average images. The data was obtained by processing 512 frames and the total image was constructed from 7 discrete frequencies.

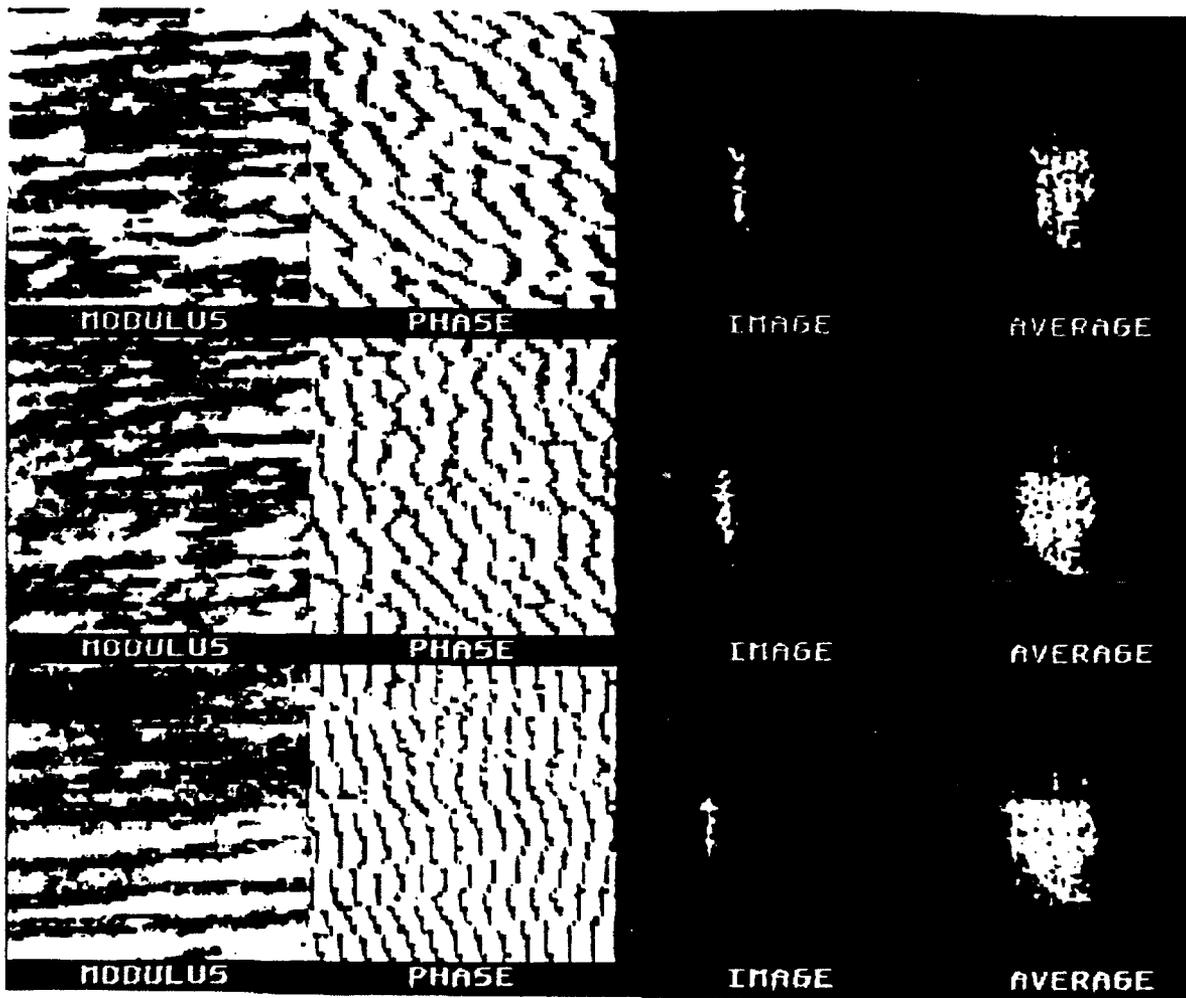


Figure 11. Last three processed frequencies from a rotating target..

Figure 11 suggests that another important effect is being observed in the data processing; averaging of a speckled image. By using the theoretical frequency bin size with equation 18, we should only be able to see a strip of the target which is 0.06 mm in width!

Clearly as seen in Figure 11 we see a far larger portion of the target at each processed bin. If closely spaced frequency bins are processed, the speckled image strips overlap as graphically depicted in Figure 12.

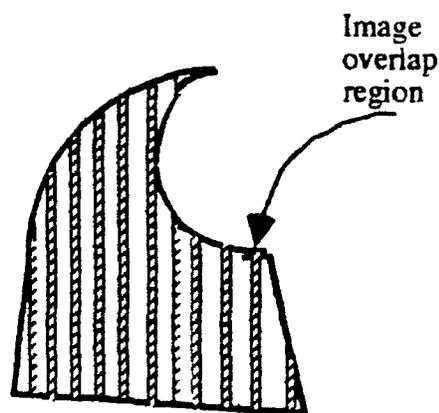


Figure 13. Overlap of processed speckled image strips.

This overlap turns out to be an advantage because the speckles are scanning thus each image strip is from a slightly different speckle perspective. When 512 frames of data are acquired the speckle perspective has changed numerous times. The speed of the transverse speckle scan can be determined by cross correlating the modulus from two consecutive 16 frame data sets. This process is outlined in the Heterodyne report from the summer of 1991.² The overall effect is the surprisingly good average images obtained by processing the maximum number of frequency bins in the Doppler broadened spectrum.

To improve on the average, the most logical step would be to average a greater number of images. But the number of images produced from the Doppler broadened data is limited by the width of the spectrum containing image information. The width of the spectrum is directly related to the velocity of the image. So, targets with larger velocities should provide more images to average. Of course equation 18 suggests that the image strips should be getting narrower and this is the case. But, as can be seen in Figure 13, the overlap region of speckle smearing is about the same for both velocities. This Figure is also good evidence of the fact that the image strip is getting smaller as the velocity increases as our equation predicted for the Doppler effect. Figure 14 display the average images obtained from targets rotating at the net radial speed give in the previous Figure. The image appears to fill in as the velocity increases, but as with all things there is a limit to the improvements.

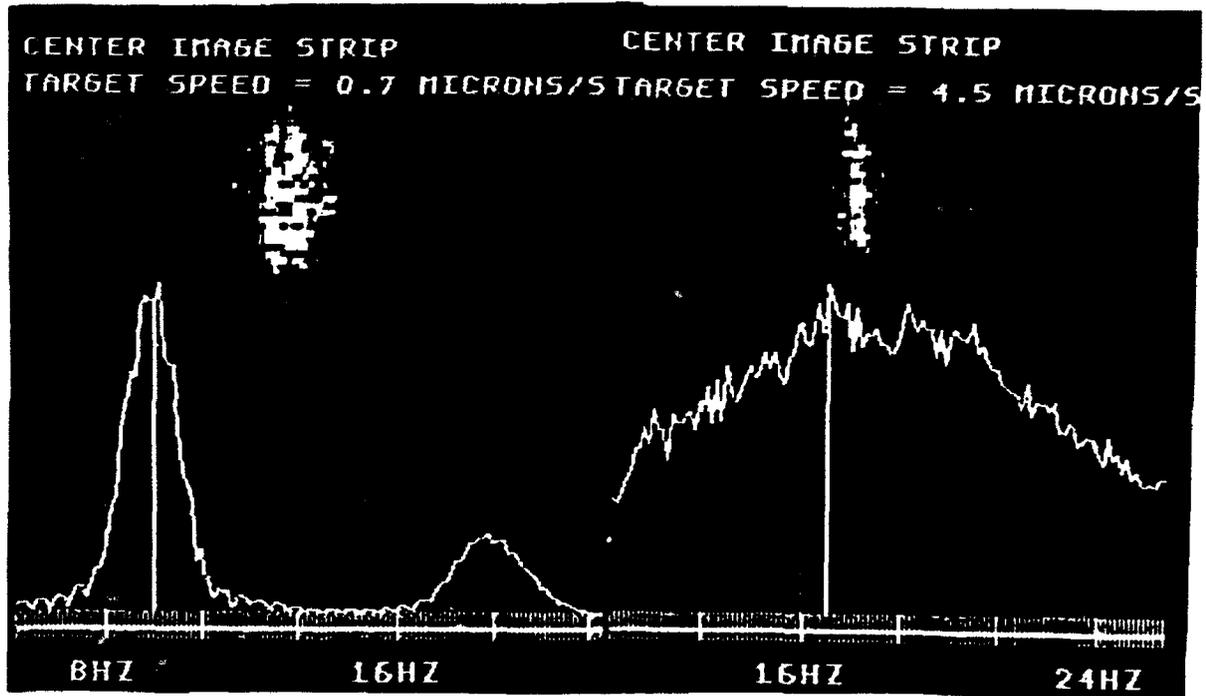


Figure 13..The effect of target speed on image width.

Note that the image in these figures has been magnified three times.

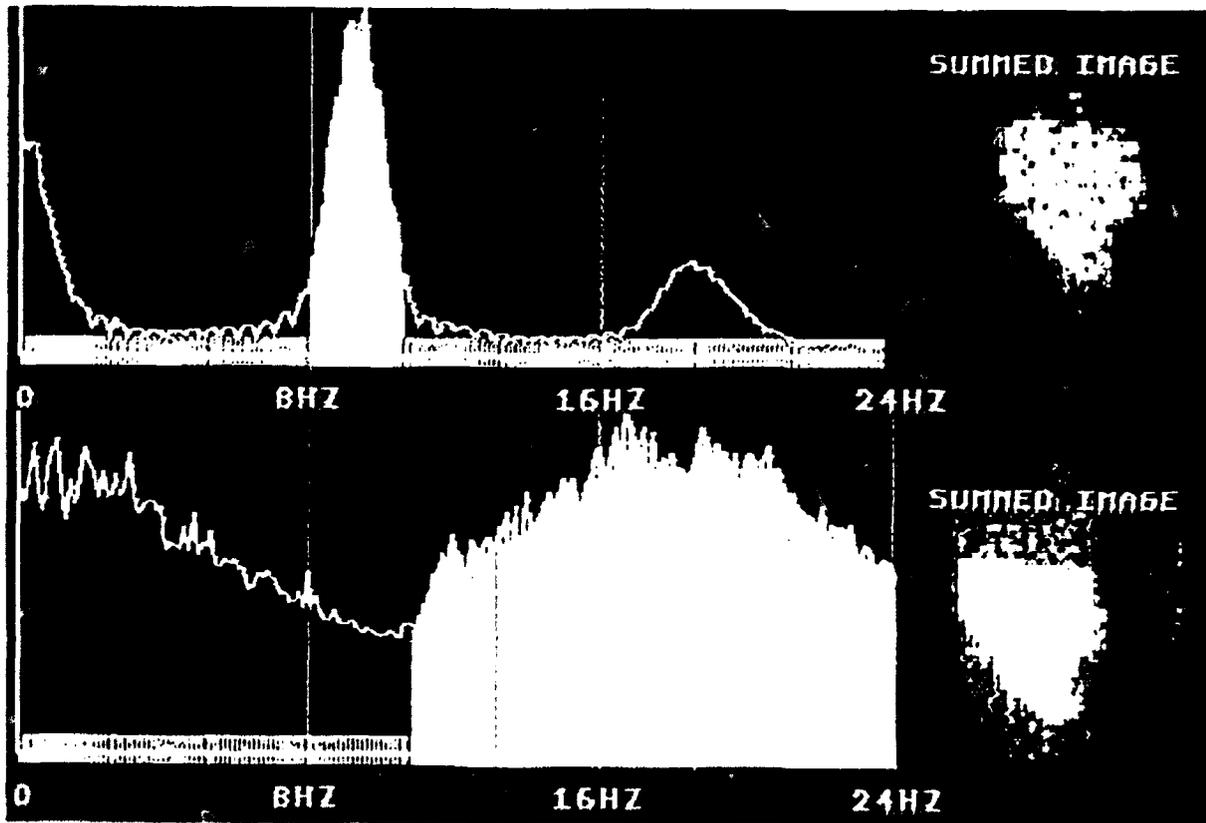


Figure 14. Averaged images at different target speeds.

The last topic to be considered in the Doppler imaging work to this date is the effect of aliasing. Clearly all experiments are bound by some effect which limits the usefulness of the process and in the case of Doppler imaging it is the effect of under sampling the higher frequencies. The primary problem is the very large frequency shifts associated with light. In this experiment, the limit on rotational motion was about 2.5 radians per second. There are two effects to consider here. The first is the effect of the entire frequency plot riding on a DC bias. This bias is the result of many higher frequencies bleeding into the data and the effect is a growing level of noise in the image. This is clearly evident in the two images of Figure 14. The second effect occurs when the width of the broadened curve begins to overlap the lower frequencies. This effect is shown in Figure 15 where three frequencies have been processed. Notice that the ghost image becomes more pronounced as the lower frequencies are processed. Also notice a definite linear phase which is not the same as the target. This suggests a possible linear phase filter. The final Figure shows the 3 hertz signal of Figure 15 along with its frequency spectrum.

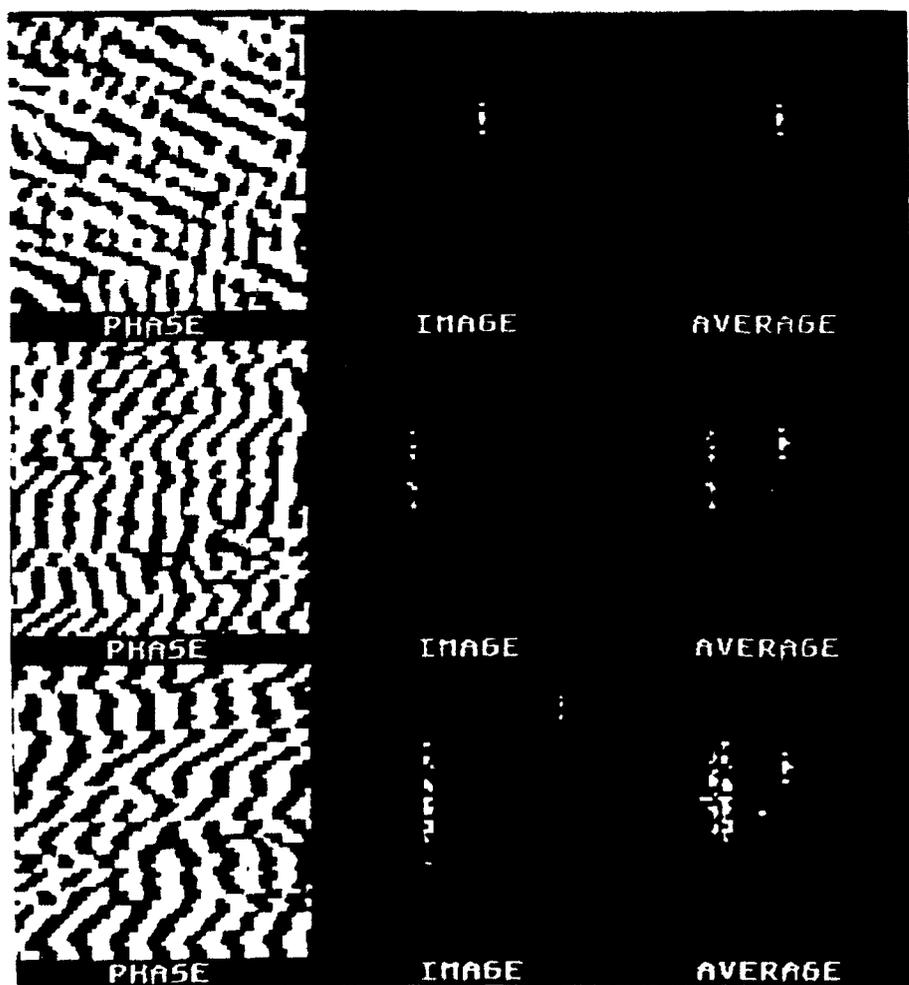


Figure 15 Three processed frequencies at 5,4 and 3 Hz.

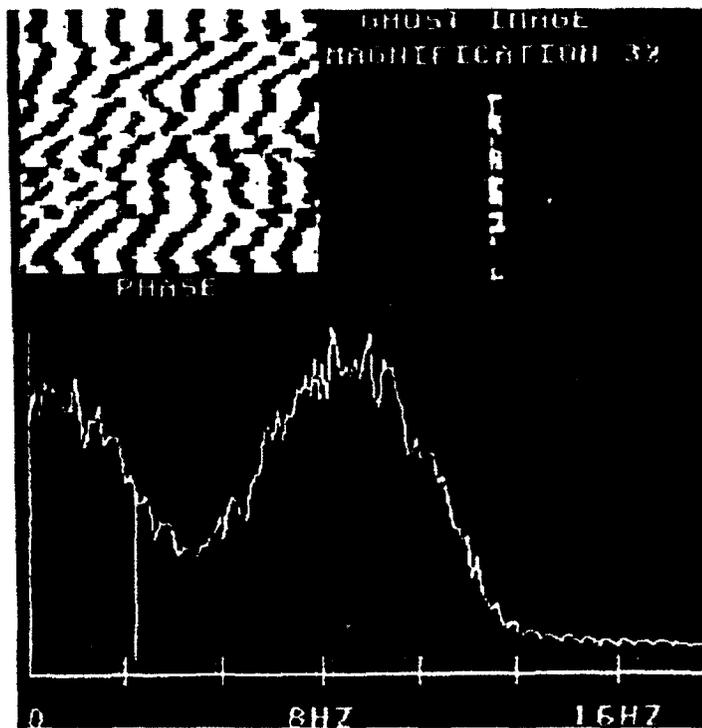


Figure 16. 3 Hz signal processed.

CONCLUSION

The Doppler imaging experiment was completely successful at wetting the researchers appetite for further investigations. Down-link phase-up could certainly be improved with time and effort. Also, this technique for phase up requires a stationary arrangement which is not the case when imaging through the dynamic atmosphere. Possibly using the algorithm present for phase-up is appropriate, but a new method for measuring distortions should be researched. The results of imaging a moving target were better than expected. While detector bandwidth may be the ultimate problem, it is not the only one. Various techniques for cleaning up the data appear to be promising such as the linear phase filter. Electronic filters at the detector array may also prove to be useful.

REFERENCES

- 1) R. Anderson, Coherent Heterodyne Array Doppler Imaging, Internal report to the Phillips laboratory, August (1992).
- 2) P.S. Idell and J.D. Gonglewski, "Image Synthesis from Wavefront Measurements of a Coherent Diffraction Field", Opt.Lett. 15, (1990).
- 3) J.K. Boger, et . al. "Experimental Techniques in Heterodyne Laser Speckle Imaging", Proc. LASERS '91 318-324, (1991).
- 4) S. E. Clark, L.R. Jones, L.F.Desandre, Coherent Array Imaging, Appl. Opt. 30, 1804-1810 (1991).

SECOND-HARMONIC GENERATION IN CORONA-POLED MATERIALS

Gene O. Carlisle
Professor
Department of Chemistry

West Texas State University
Canyon, Texas 79016

Final Report for:
Summer Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

September 1992

SECOND-HARMONIC GENERATION IN CORONA-POLED MATERIALS

Gene O. Carlisle
Professor
Department of Chemistry
West Texas State University

Abstract

Second-harmonic generation (SHG) and spectroscopic absorption measurements were used to study the nonlinear optical (NLO) properties of three polymeric thin ($\approx 1 \mu\text{m}$ thick) films and fused quartz. These materials were made optically nonlinear by a parallel-wire corona-poling procedure. Two of the polymers were side-chains containing NLO chromophores covalently attached to every other carbon atom in acrylate chains while the third polymer was a guest-host system of azo-dye guest molecules in the host poly(methyl methacrylate). For the first time, commercial grade fused quartz, 3 in x 1 in x 1 mm, was made optically nonlinear by a parallel-wire corona-poling procedure. The orientational order parameters of the polymeric thin films were determined from polarized absorption spectra; these stabilized parameters, ranging as high as 0.51, are higher than any reported in the literature. Second-harmonic intensities were measured by the Maker-fringe technique. Because of the very high number density, the side-chain polymers were considerably more efficient in producing SHG than the guest-host films or the fused quartz.

SECOND-HARMONIC GENERATION IN CORONA-POLED MATERIALS

Gene O. Carlisle

INTRODUCTION

The use of optics in electronic devices for increased speed and capacity has been limited, at least in part, by the few good nonlinear materials. For electro-optic modulators, materials must possess not only a high second order susceptibility, but also such attributes as ease of preparation, fabrication into waveguides, and stability. Present experimental devices are based on lithium niobate because it is the best available among a rather unsatisfactory set of inorganic materials.

The use of organic materials [1,2] for nonlinear optical (NLO) processes has been gaining increased attention because these materials offer a number of advantages including a large variety of molecular structures with large nonlinearities and the capability for designing other desirable properties.

The growth of noncentrosymmetric crystals, preparation of Langmuir-Blodgett films, and the poling of charge transfer molecular groups in polymeric matrices are the basic approaches being taken to develop organic NLO materials. Because of the problems [3] with the first two, the research at Phillips Laboratory, PL/LITN, focussed on the latter. Corona poling [4] was chosen as the method of poling since it generates higher electric fields before breakdown as compared to flat-electrode poling.

Second-harmonic generation (SHG) and spectroscopic absorption measurements were used to study the nonlinear optical (NLO) properties of three polymeric thin

($\approx 1 \mu\text{m}$ thick) films and fused quartz. These materials were made optically nonlinear by a parallel-wire corona-poling procedure. Two of the polymers were side-chains containing NLO chromophores covalently attached to every other carbon atom in acrylate chains while the third polymer was a guest-host system of azo-dye guest molecules in the host poly(methyl methacrylate). For the first time, commercial grade fused quartz, 3 in x 1 in x 1 mm, was made optically nonlinear by a parallel-wire corona-poling procedure.

PREPARATION OF MATERIALS

DR1-PMMA. The 95% pure DR1 obtained from Aldrich was recrystallized from a hot benzene solution and dried overnight in a vacuum oven at 68° C. The structural formulas and synthetic method are shown in Fig. 1. 0.0055 mole of recrystallized DR1 (1) was dissolved in 80 mL of benzene at 75° C, followed by equal molar amounts of the base triethylamine and methacryloyl chloride. The reaction mixture was allowed to stir over night at room temperature. After the precipitated triethylammonium chloride was removed by filtration, the solvent was removed by rotary evaporation, and the solid azo dye substituted acrylic monomer (2) was recrystallized from an ethanol:acetone (120 mL:20 mL) mixture. The maroon crystals were dried in a vacuum oven at 68 degrees for 3 hrs. The yield was 61% and the melting point was 81° C. Elemental analyses for carbon, hydrogen, and nitrogen were in excellent agreement with the theoretical values for structure (2).

Polymerization of the monomer (0.5357 g) was carried out in the presence of the initiator azoisobutyronitrile (AIBN) dissolved in a toluene:ethanol (2 mL:1 mL) mixture and under a 30 psi nitrogen atmosphere at 75° C. The polymer

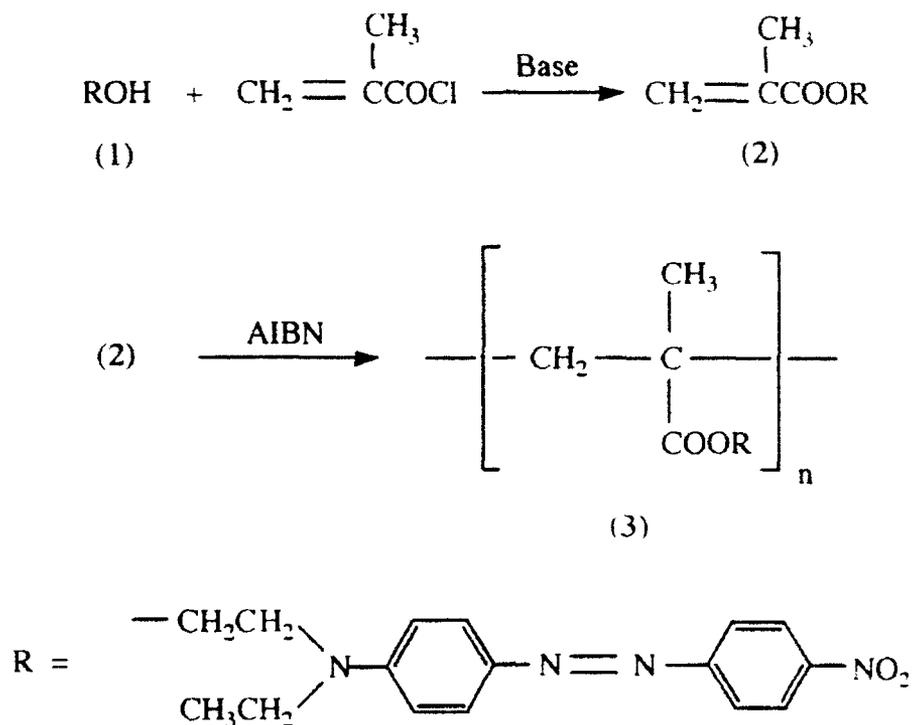


FIG. 1 Structural formulas and synthetic method for DR1-PMMA.

(3). DR1-PMMA, was collected, washed with acetone, and further purified by precipitation from a cyclohexanone solution. Finally, a 40% yield of the polymer was obtained after drying in a vacuum oven at 75⁰ C for 24 hours. Elemental analyses for the polymer were also in excellent agreement with structure (3). IR spectra obtained from KBr disks of the monomer and polymer confirmed the structures as shown in Fig. 1.

The films were prepared from solutions containing 0.1047 g of polymer dissolved in 0.80 mL of cyclohexanone and 0.20 mL of nitrobenzene. The mixture was heated at 90⁰ C while stirring for one hour to ensure that a homogeneous solution was present. The solution, while still hot, was spun coated (Headway

Research spinner) at 1640 rpm for one minute onto 3 in x 1 in x 1 mm thick transparent microscope slides (Fisher Premium). The films were then baked in a vacuum oven at 90⁰ C for 2 hrs to remove residual solvents. The resulting films were found by the use of a profilometer (Talystep) to have a typical thickness of 0.344 μm . The surface of the films proved to be resistant to minor scratches as the stylus of the profiler did not show any tracks on the films even when viewed at a magnification of 250. The density of the polymer was found to be 1.19 g/cm^3 , and the concentration of the DR1 chromophore was determined to be $1.87 \times 10^{21}/\text{cm}^3$.

The films were corona poled with a configuration and procedure similar to those of Mortazavi, *et al* [4] as shown in Fig.2. The substrate was placed, with

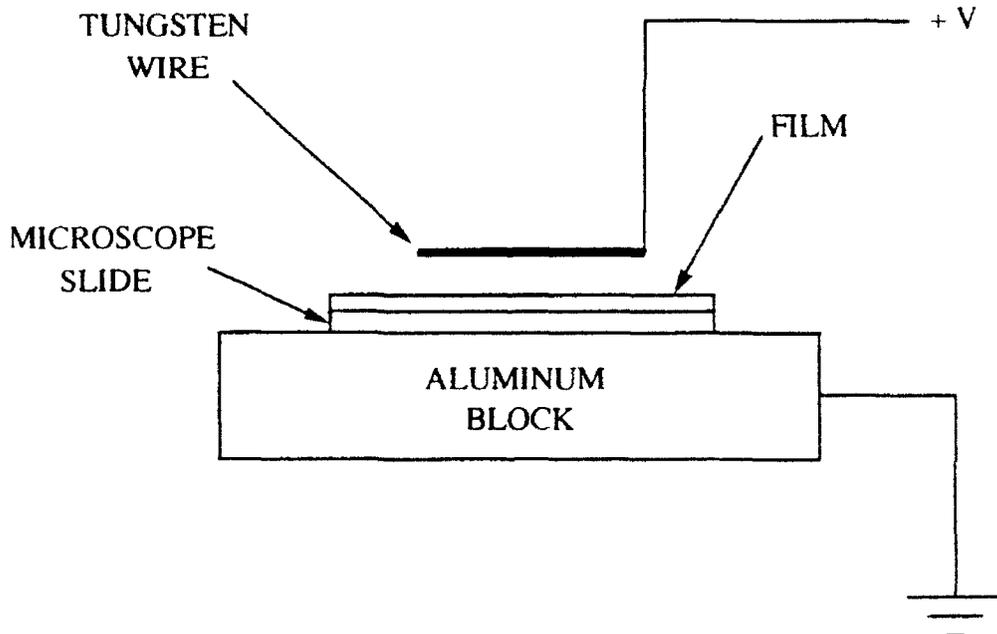


FIG. 2 Diagram of corona poling.

the film side up, on the center of a grounded planar aluminum electrode (4 x 4 x 1 in). The temperature of the film could be controlled by an appropriate setting of a hot plate positioned under the aluminum electrode. The actual temperature of the film could be measured by an iron-constantan thermocouple wire pressed against the film. A 25- μ m diameter and 4.0-cm length tungsten wire (Goodfellow) was held parallel to and 4.0 mm above the film. The glass-rubber transition temperature was measured with a differential scanning calorimeter (Shimadzu DSC-50) and found to be 128.5⁰ C. Therefore, in order to ensure good mobility of the chromophore units during poling, the temperature of the films was raised to 131⁰ C. A positive voltage was then applied to the wire electrode, and the current from the planar electrode was monitored by measuring the voltage drop across a 10-M Ω resistor in series. The voltage to the wire electrode was slowly increased with caution until a corona breakdown voltage of 3.2 kV was reached which produced a current of 3.9 μ A. After poling for 30 min, the hot plate was turned off and the voltage increased, as the temperature of the film decreased, in order to maintain a current of 3.9 μ A until the temperature reached approximately 85⁰ C. The voltage at this point was 5.0 kV which was then maintained until the film reached room temperature. The entire poling procedure usually required about 3 hours. The procedure consistently produced films of high optical quality. The poling area, which appears much more transparent than the unpoled film, is approximately 2 cm wide and extends the length of the wire electrode.

Polarized absorption spectra were measured for the films before and after poling with a Shimadzu 2100U spectrophotometer equipped with Glan Taylor polarizers (Oriel). The spectra were recorded over the range of 300 - 700 nm.

More details on the preparation, poling, and absorption studies of DR1-PMMA may be found in our publication [5].

PCMA. The structure of poly(coumaromethacrylate) (PCMA) is shown in Fig. 3. Thin films, 3.5 μm , of (PCMA) were prepared, corona-poled, and their spectra

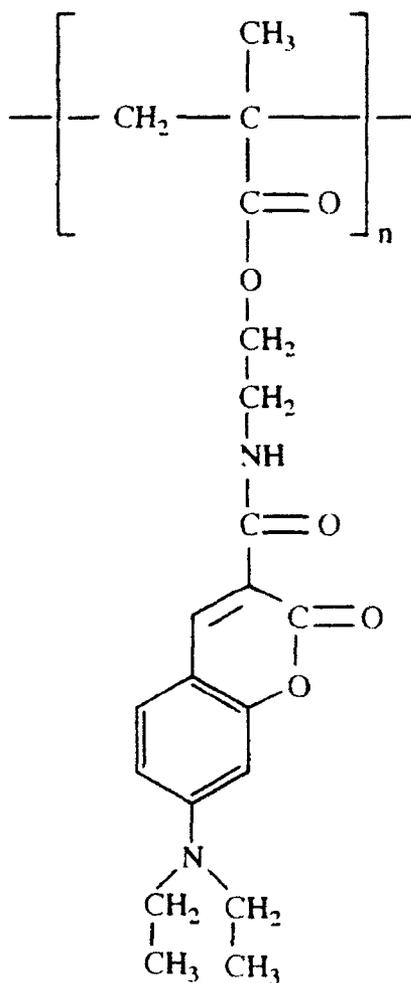


FIG. 3 Chemical structure of PCMA.

measured in a similar manner to that of DR1-PMMA as described above.

DO3/PMMA. The structures for this guest-host system of disperse orange 3 (DO3) and poly(methyl methacrylate) (PMMA) are shown in Fig. 4. The films were

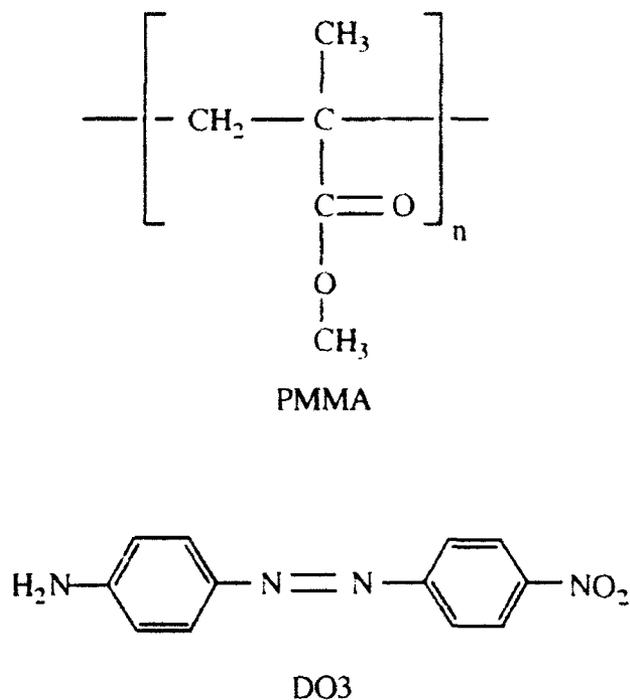


FIG. 4 Chemical structures of PMMA and DO3.

prepared by a procedure similar to that described above for the DR1-PMMA films. The resulting films containing 10% DO3 were found to have thicknesses in the range of 1 μm . The concentration of DO3 in the films is about 1.5×10^{20} molecules/cm³.

The films were corona poled with a configuration and procedure similar to that described above. The glass-rubber transition temperature was found to be 81 C. In order to ensure good mobility of the guest molecules, the temperature of the films was raised to 85⁰ C for poling. Polarized absorption spectra were

measured for the films as described above. More details on the preparation, poling, and absorption studies of these DO3/PMMA films may be found in our publication [6].

Fused Quartz. 3 in x 1 in x 1 mm slides of T08 Commercial fused quartz were obtained from Heraeus Quartz. The slides were corona-poled as described above using a 4-cm length, 50- μ m diameter tungsten wire held parallel to and 9.0 mm above the quartz slide. Typical poling conditions were: 280⁰ C poling temperature, 20 minutes poling time, 8.0 μ A current, and 9.0 kV applied voltage.

SECOND-HARMONIC GENERATION MEASUREMENTS

SHG measurements were made using the experimental setup shown in Fig. 5. A Lumonics Nd:YAG laser, wavelength 1.064 μ m, Q-switched at 10-Hz repetition rate, and 150 mJ per 15-ns pulse was used. The p-polarized beam from the laser passes through the IR-pass filter (F1). PBS1 is a polarizing beam splitter allowing only p-polarized light to pass on to M1, a p-reflecting mirror. After going through the aperture (A) of 5.0 mm diameter, the unfocussed beam is split by M3, a mirror designed to reflect 30% and pass 70% of the p-polarized beam. The reflected beam passes through the reference (R), which was a quartz crystal, and then to M4, a mirror designed to reflect the IR and pass the second harmonic (532 nm) generated by the quartz reference. The green light passes through the green-pass filter (F3) and the lens (L). It is then reflected from a beam stirring mirror (M5) and passes through two interference filters (IFs) before being focussed on the photodiode detector. The plane polarization of the IR beam passing through M3 can be varied as it passes through the half-wave plate (HW).

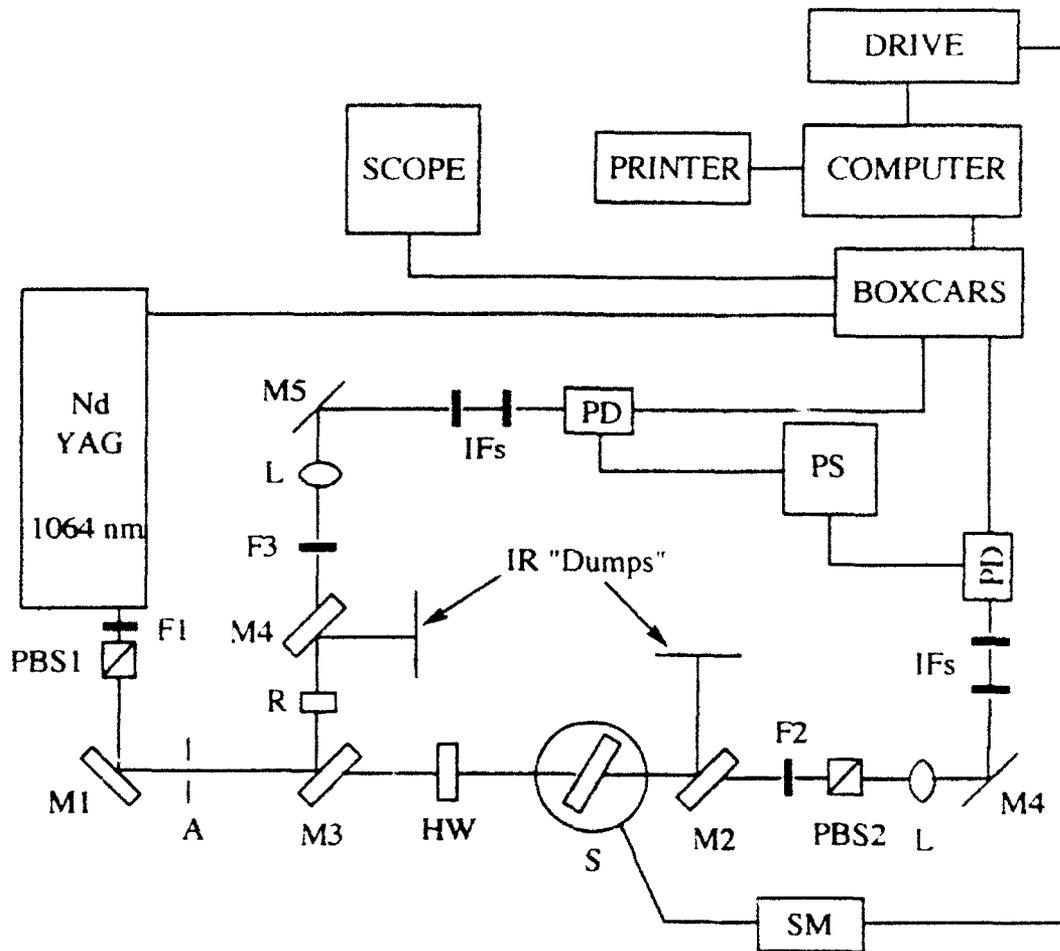


FIG. 5. Experimental setup for SHG measurements.

Experiments using p-polarized, s-polarized, and 45° -polarized fundamentals were used to excite samples (S) positioned on a rotation stage. M2 reflects the IR and passes the second-harmonic light generated by the sample. After passing through the green-pass filter (F2), a polarizing beam splitting cube is used to ensure a specific polarization of the second-harmonic to be measured. After passing through the lens and reflected from M4, the beam passes through the interference filters and on to the photodiode detector. Neutral density filters were added at

the F2-position if required to avoid saturation of the sample diode. Experiments using p-excited p-detected, s-excited p-detected, and 45^o-excited s-detected were performed. Signals from the sample and reference detectors are sent to boxcar averagers and then stored the computer. The computer also controlled the stepping motor such that SHG intensities from the sample could be measured as a function of angle of incidence. Data were recorded from -80^o to +80^o.

RESULTS AND DISCUSSION

Results and discussions regarding preparation, corona poling, absorption studies, calculations, and theoretical models for DR1-PMMA and DO3/PMMA films may be found in our publications [5,6]. Order parameters for the molecular dipoles were calculated from the absorption data using a rigid oriented gas model. Stabilized order parameters of 0.51 and 0.30 were found for DR1-PMMA and DO3/PMMA films, respectively. These parameters, which are the highest reported for side-chain and guest-host systems, have remained constant for over 18 months. A stabilized order parameter of 0.17 was found for PCMA films. The lower order parameter for the PCMA films is very likely due to the lower dipole moment of the side-chain chromophore which results in less poling torque by the field. These films have remained stable for over two years.

A very large amount of SHG data, over 600 data files, were taken on the three polymeric materials and fused quartz samples. The SHG intensity vs angle of incidence is shown in Fig. 6 for a DR1-PMMA film. For each material similar Maker fringes were observed. The results showed a high degree of poling for each polymeric film as intensities were several times greater than that of the quartz reference. The SHG intensities of the DR1-PMMA and PCMA films were about the

same: however, since the PCMA films were about 10 times thicker (3.5 μm vs 0.34 μm), the second-order nonlinear susceptibility, $\chi^{(2)}$ should be considerably higher for the DR1-PMMA films. This is due to the higher order parameter and higher β value for the NLO chromophores of the DR1-PMMA polymer. The guest-host DO3/PMMA films with an order factor of 0.30 and a number density of about 1/10 that of DR1-PMMA and PCMA exhibited SHG intensities much less than those of these side-chain polymers. In order to calculate the important elements of the $\chi^{(2)}$ tensor, applications of the appropriate theories to the SHG data are now in

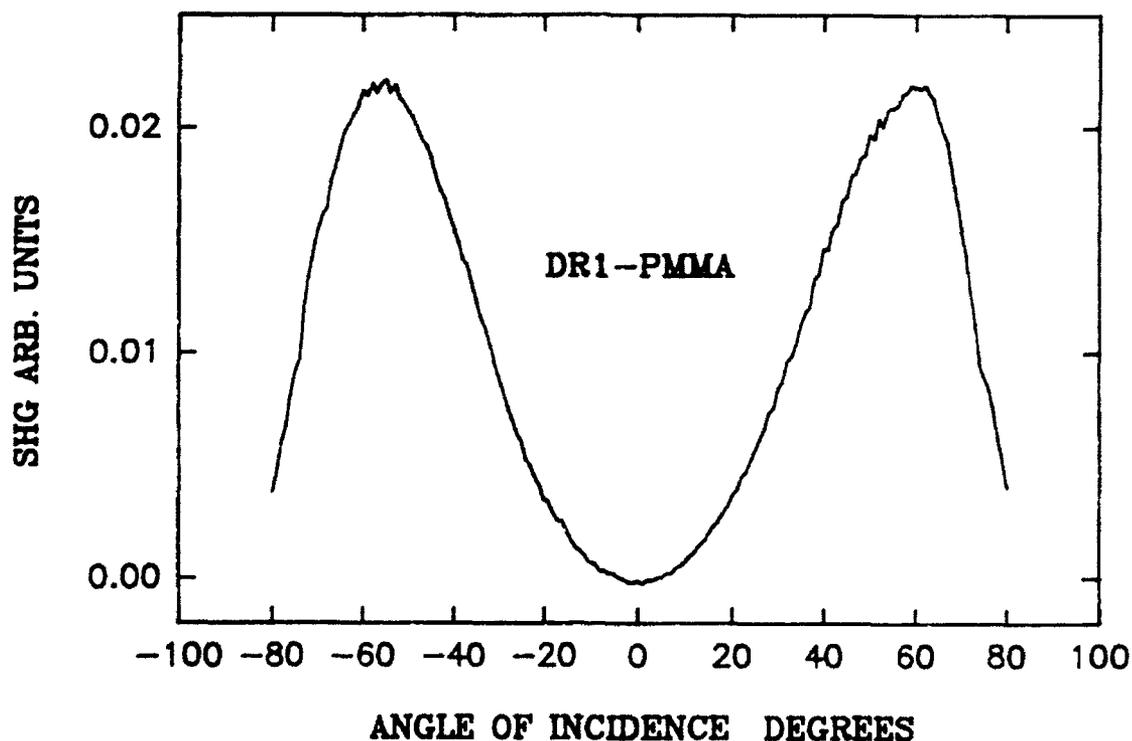


FIG. 6. SHG vs angle of incidence for DR1-PMMA.

progress.

SHG intensities produced by the corona-poled fused quartz samples were considerably less than those for the polymers. This is likely due to a much lower β value for the localized moieties in the quartz compared to those of the organic chromophores in the polymers. Also, the results show that the quartz is not poled throughout its entire thickness of 1.00 mm. Through controlled etching experiments with HF solutions, poled quartz samples were etched to various depths. Subsequent profilometer and SHG measurements were then made, and the poling depths were found to be only about 4 μm . Fused quartz samples were also flat-electrode poled using a stainless steel electrode. SHG intensities from corona-poled samples were significantly higher than those of flat-electrode poled samples. Although the exact mechanism by which poled quartz produces SHG is not established, it is believed that impurities play an important role [7]. Accurate chemical analyses using atomic absorption methods are now in progress to determine the exact percentages of aluminum and other impurity elements possibly responsible for the NLO properties of poled fused quartz.

CONCLUSION

SHG and absorption data were measured for three polymeric materials and fused quartz. Data are essentially complete and are now being analyzed in terms to the appropriate theories. The two side-chain polymers, DR1-PMMA and PCMA, are expected to exhibit very high $\chi^{(2)}$ values. For the first time quartz slides have been made optically nonlinear by corona poling with a parallel wire configuration. SHG intensities were significantly higher for corona poled than for flat-electrode poled quartz. Two papers were published on the preparation.

poling, and absorption studies of the DR1-PMMA [5] and DO3/PMMA [6] films. A paper on the DR1-PMMA films will be presented at the 48th Southwest Regional Meeting of the American Chemical Society in Lubbock, Texas on 21 October 1992. Four more papers regarding SHG studies are expected to be published on these materials within a year. Further details of this research are described in my laboratory notebooks and in computer files at Phillips Laboratory (PL/LITN).

REFERENCES

1. S. R. Marder, J. E. Sohn and G. D. Stucky, (editors) *Materials for Nonlinear Optics* (American Chemical Society, Washington, DC, 1991).
2. P. N. Prasad and D. J. Williams, *Introduction to Nonlinear Optical Effects in Molecules and Polymers* (Wiley, New York, 1991).
3. M. Eich, A. Sen, H. Looser, G. C. Bjorklund, J. D. Swalen, R. Twieg, and D. Y. Yoon, *J. of Appl. Phys.* 66, 2559 (1989).
4. M. A. Mortazavi, A. Knoesen, S. T. Kowel, B. G. Higgins, and A. Dienes. *J. Opt. Soc. of Am. B* 6, 733 (1989).
5. F. K. Ratsavong, G. O. Carlisle, and D. R. Martinez, *J. Mat. Sci. Mat. Elect.* 3, 187 (1992).
6. S. Hu, G. O. Carlisle, and D. R. Martinez, *J. Mat. Sci. Let.* 11, 794 (1992).
7. R. A. Myers, N. Mukherjee, and S. R. J. Brueck, *Opt. Let.* 16, 1732 (1991).

ACKNOWLEDGMENTS

I thank Dr. Christopher M. Clayton and Mr. David R. Martinez for the opportunity of returning for a second summer of research at the NLO Center of Technology. The enjoyable working atmosphere generated by them. Major Lynn M.

Black, Dr. Stephen H. Chakmakjian, Captain Victor C. Esch, Dr. Mark T. Gruneisen, Ms. Gladys J. Kassay, Dr. Karl W. Koch, Dr. Mark W. Kramer, Dr. Gerald T. Moore, Dr. Phillip R. Peterson, Dr. Kenneth D. Shaw, Dr. P. Mohinder Sharma, and Captain Karen A. Yackiel is very much appreciated. The research accomplished this summer is greatly attributed to considerable efforts by Mr. Martinez and Dr. Koch.

The sponsorship of the Air Force Office of Scientific Research and Phillips Laboratory is gratefully acknowledged. The administrative support of Research and Development Laboratories is much appreciated.

USE OF OPTICAL FIBERS
IN LONG BASELINE INTERFEROMETRIC IMAGING

Douglas A. Christensen
Professor
Department of Electrical Engineering

University of Utah
Salt Lake City, Utah 84112

Final Report for:
AFOSR Summer Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington D.C.

September 1992

USE OF OPTICAL FIBERS IN LONG BASELINE INTERFEROMETRIC IMAGING

Douglas A. Christensen
Professor
Department of Electrical Engineering
University of Utah

Abstract

Optical fibers offer the promise of flexibility, ease of use, and low cost when used to recombine the light in the two branches of a long baseline interferometer. Such interferometers are being studied for imaging stellar objects with high resolution, including geosynchronous satellites, the sun, and stars. The work described in this report is a continuation of a study to ascertain the parameters that are important in implementing fibers in a broadband interferometer. The factors which must be considered include dispersion effects in the fibers, phase variation or wander in each of the branches and how to control the phase, and the low light levels which will accompany the use of single mode fibers. Toward these goals, two tasks were achieved this summer: 1) Broadband interference fringes were found in an intermediate-step interferometric setup in the laboratory, namely a side-by-side arrangement with two subaperture telescopes observing the same quasi-point source and feeding two 60 m long fiber paths before recombination; and 2) A phase-lock condition was obtained in a test fiber interferometer using a piezoelectric fiber modulator and a lock-in amplifier system. This "internal metric" phase-locking condition is important for the eventual application of the interferometer for imaging purposes.

USE OF OPTICAL FIBERS IN LONG BASELINE INTERFEROMETRIC IMAGING

Douglas A. Christensen

INTRODUCTION

This summer's work was part of a larger project whose overall goal is to develop a fiberoptic-linked long baseline interferometer for imaging stellar objects such as geosynchronous satellites and stars. The accomplishments for the summer fell into two areas:

1. Demonstrating broadband (10 nm) fringes in progressively more realistic interferometer configurations. Last summer, broadband fringes were found in a Mach-Zehnder interferometer arrangement that included 60 m long fiber branches and which used beamsplitter cubes. The short coherence length associated with a 10 nm linewidth source requires that the optical path differences (OPD) in the two fibers be matched to within approximately 200 μm . This was accomplished using a unique feature of a laser diode: its coherence length can be adjusted by changing its drive current.

This summer, more complex configurations were set up and broadband fringes detected by matching the path lengths with the aid of a laser diode. The first arrangement employed two 8" aperture telescopes facing each other to collect the light from a large planar beamsplitter. Following successful attainment of broadband fringes, a more realistic configuration was set up in which the two telescopes were rotated into a side-by-side orientation to observe a single distant (27 feet) quasi-point source. Again, broadband fringes were obtained.

2. Achieving a phase-locked condition with a fiberoptic Mach-Zehnder interferometer incorporating a piezoelectric stress phase modulator. This condition is essential for the eventual operation of a long baseline interferometer imager. In such a device, the detected fringe shift comes from scanning the distant object under observation, and local fringe shifts due to ambient changes in the OPD (from local temperature changes or vibration, for example) must be cancelled out. This is the role of the "internal metric" procedure which monitors and cancels by feedback any changes in phase within the interferometer itself. Successful demonstration of phase locking in the interferometer was thus a valuable stepping stone to eventual implementation of a full internal metric capability.

BROADBAND INTERFEROMETRIC CONFIGURATIONS

Face-to-Face Telescope Arrangement Using Meade 8" Telescopes -

When this summer's work began, the arrangement shown in Fig. 1 was already set up on the optics table. Figure 1 shows a fiberoptic interferometer using two 8" Meade telescopes receiving light from a large planar beamsplitter. Fringes had been detected using the green HeNe laser with its long coherence length. It remained to be determined whether broadband fringes from the xenon lamp could be found. As an OPD-matching aid--as mentioned in the Introduction--the laser diode was substituted as source. With large drive current, its coherence length is quite long (tens of centimeters, almost as long as the HeNe laser) and fringes were easy to find. When the drive current is reduced, the coherence length gets considerably shorter (and output power drops), finally contracting to that expected for a broadband source when the laser diode is emitting incoherent spontaneous radiation only. By continually adjusting the air path in one branch (and therefore the OPD) to maximize the fringe contrast as drive current is reduced, one can move in a direction of a zero OPD condition.

This technique allowed the detection on laser diode fringes, but they were weak at low drive current due to two causes:

1) The fiber polarization axes were not aligned in the same direction in each of the branches. The fibers used are York "Hi Bi" bowtie stress induced birefringent single mode fibers; they will preserve polarization with respect to the orientation of their bowties. For maximum fringe contrast, the bowtie orientation of the two inputs should be matched, as well as at the two output ends. As it stood, the introduction of linearly polarized light at the input to each fiber branch resulted in elliptically polarized light at the outputs. This meant that the bowties were neither horizontal or vertical at the input end. Realignment was performed until all four fiber ends had horizontal bowtie orientation. This step increased fringe contrast noticeably.

2) The power detected out of Fiber B was always much lower (by a factor of 4X, as estimated by eye) than that from Fiber A, in spite of prolonged and repeated efforts to maximize the coupling alignment into the output end of Fiber B. It was surmised, and eventually confirmed by direct microscopic inspection, that some damage had occurred to the face of Fiber B. A microscope (magnification 100X, front light illumination looking perpendicular to the face) showed a groove or scratch that appeared to cross the central core region of Fiber B; in contrast, Fiber A showed a smooth, clear face. This is most likely the cause of the lower power in Fiber B.

DAC 5/15/92

FIBEROPTIC INTERFEROMETER w/ FACING TELESCOPES

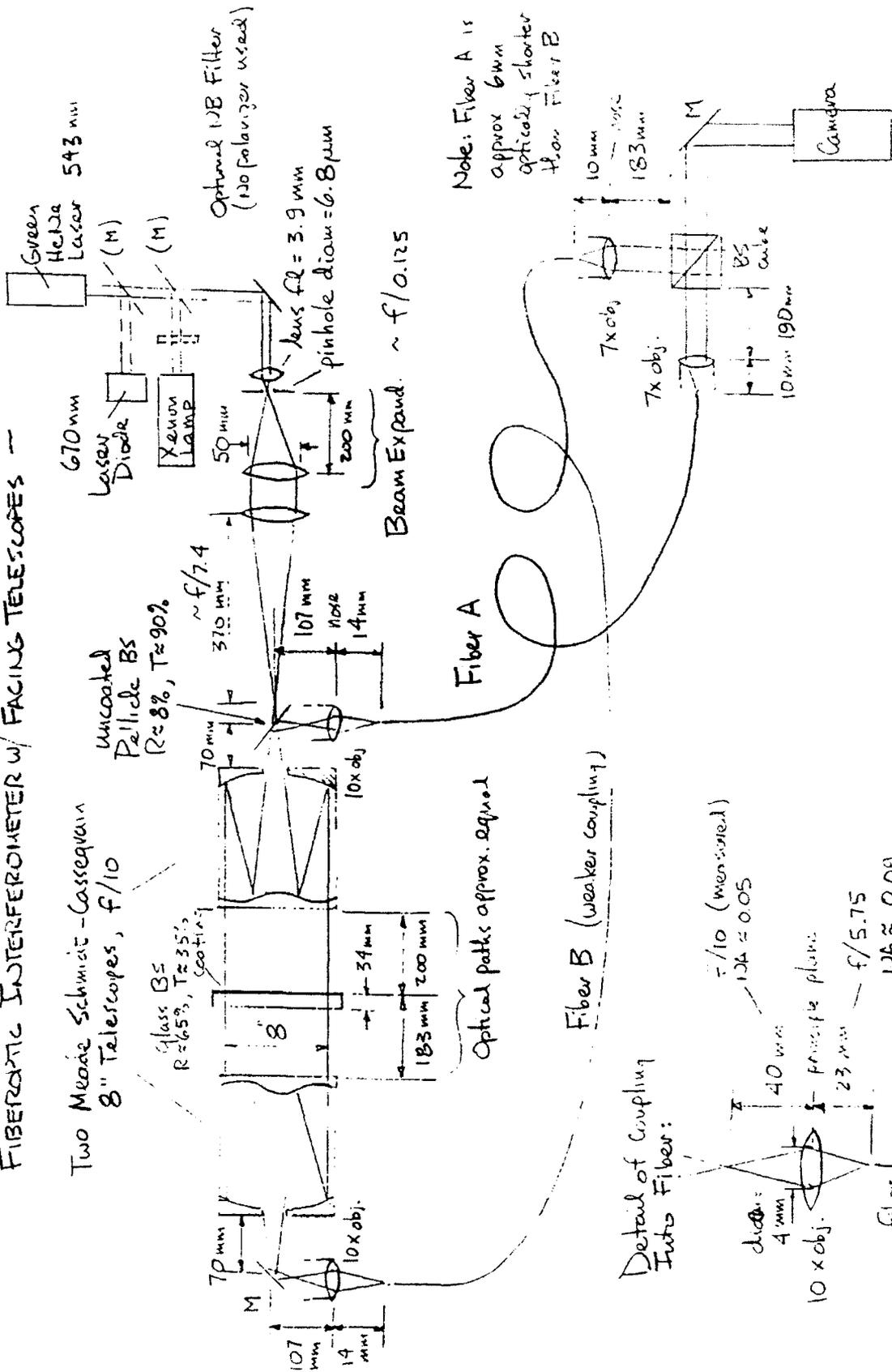


Fig. 1 - Fiber optic interferometer using two 8" Meade telescopes receiving light from a large planar beamsplitter.

To correct the problem, three remedies may be considered:

a) Try polishing the fiber face while still in the connector. This is simple and may work if the scratch is not too deep.

b) Remove the connector, cleave the fiber, epoxy on a new connector, and polish the end. This is time consuming and would require an identical treatment of the fiber on the other spool to maintain length equality. It is a last choice alternative.

c) Put index-matching gel on the fiber face (to match away the scratch somewhat), then place a small piece of glass cover slip on the face, held in place by surface tension, to give a planar surface. This is, of course, a temporary solution, but is easy to do.

For now, rather than trying any of these three ideas, the power in the two branches was approximately equalized by simply replacing the previous coated 40/40 pellicle beamsplitter, near the microscope objective which couples into Fiber B with an uncoated 8/90 pellicle beamsplitter, then switching Fiber A for Fiber B. This trick worked in equalizing the light power, and, after adjusting the OPD with the laser diode, white light fringes were finally obtained.

One observation became evident while watching these fringes for a few minutes: Any airborne or acoustic sounds (talking, clapping, doors closing, etc.) caused dramatic fringe instabilities in the output pattern. This could be reduced somewhat by closing all the doors in the table cover, but even then it was noticeable. The phenomenon was eventually traced to the pellicle beamsplitter which was acting as a very sensitive microphone, vibrating when sound waves passed through it. Solutions would be to: a) completely acoustically isolate the beamsplitter, or b) replace it with a solid glass device. Since we were soon going to the next configuration, neither solution was felt needed.

Side-by-Side Telescope Arrangement -

The next step toward a progressively more realistic setup is shown in Fig. 2. This figure shows a fiberoptic interferometer in which light is collected from a single common source by side-by-side Meade 8" telescopes. Since the closest object distance that the Meade telescopes can focus to is between approximately 35 to 50 feet, the challenge was to position the (apparent) source this far away from the telescopes. This could be done by moving the source out into the hallway adjacent to the lab, but it would be more convenient to stay completely on the optical bench at this stage. A solution was found by folding the optical path from source to telescopes on the bench, as

DAC 5/28/92

FIBEROPTIC INTERFEROMETER w/ FOLDED PATH

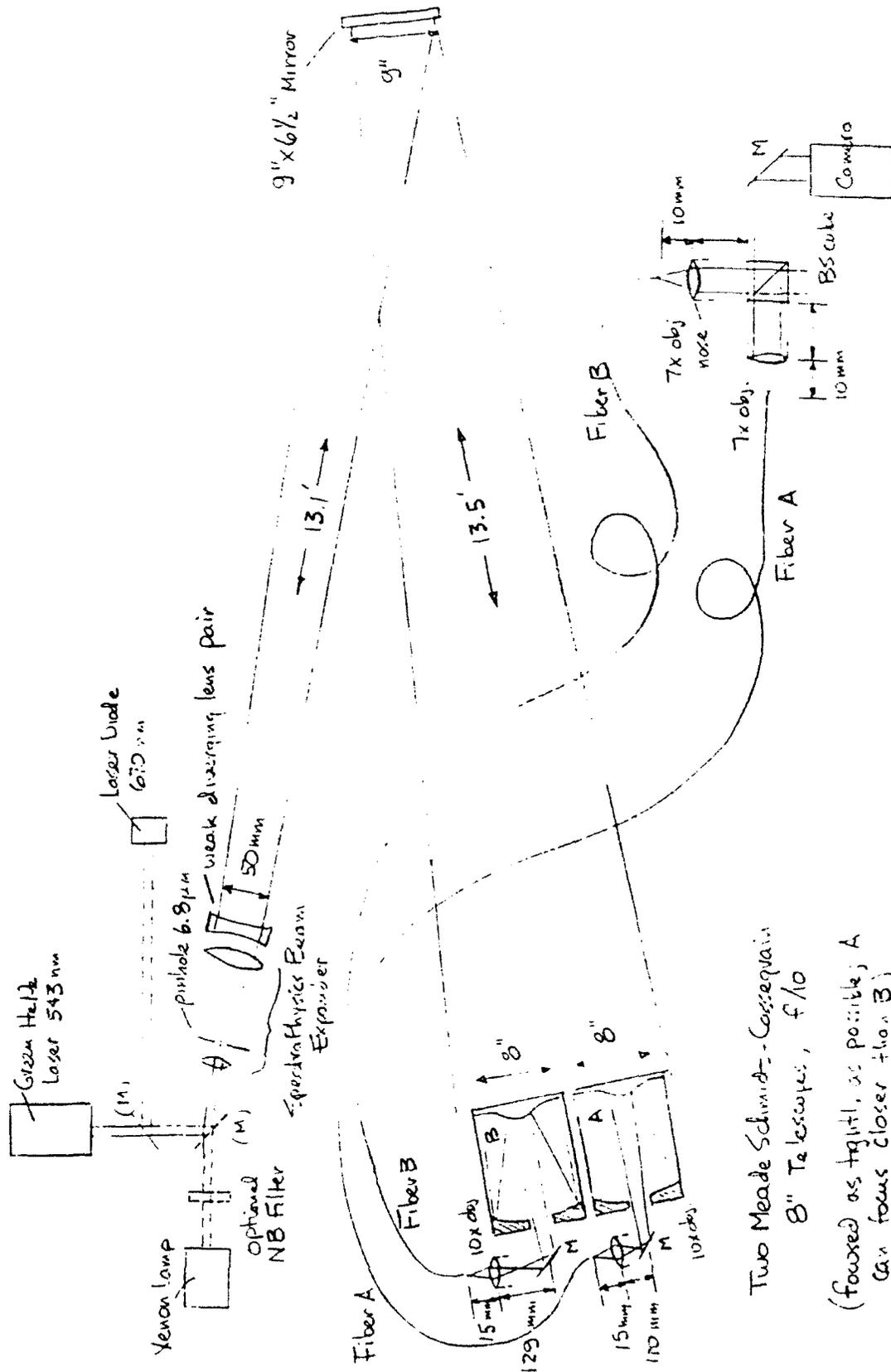


Fig. 2 - A fiber optic interferometer in which light is collected from a single common source by side-to-side Meade 8" telescopes.

diagrammed in Fig. 2.

While setting up this new configuration, efforts were made to match the Numerical Aperture (NA) of the light exiting the 10X microscope objectives to the acceptance NA of the fibers; from York's specifications, this is $NA = 0.14$. This was accomplished by calculation and subsequently positioning of the fiber faces with respect to microscope objective lenses, with the values shown in Fig. 2.

When the new setup was completed, light from the green HeNe laser source was coupled into the interferometer. Again, the power in Fiber B was much lower than in Fiber A, but since no beamsplitters are employed in this configuration, power balancing could not be achieved by altering the beamsplitting ratio. Rather, the end of Fiber B was gelled and a small piece of cover slip was affixed to the fiber face. This improved the power in Fiber B, and, although it still was not as much as in Fiber A (by a factor of 2-4X), it did allow good HeNe fringes to be seen. Note that the distance measurements shown in Fig. 2 indicate that Fiber A is about $(166 + 110) - (150 + 129) = 3$ mm optically longer than Fiber B. This is in qualitative agreement with measurements from last year showing that Fiber A is 6 mm optically longer than Fiber B. Note that the most recent values are more susceptible to measurement error since two covered telescopes are now in the path; thus the 6 mm value is probably the more accurate number.

Upon observing these fringes for a period of time, it became clear that this setup was much more susceptible to amplitude variations due to air movement and vibration than the previous (face-to-face) arrangement. These instabilities were not phase variations (such as caused by the previous pellicle beamsplitter, for example) but rather were amplitude instabilities caused by the focused spot on the fiber faces moving off the core. The larger effective "optical lever" inherent in this configuration is probably the cause of the increased sensitivity to vibration. An active beam positioning device, prototyped by OPTRA and refined by Rockwell Power Systems, is planned to be tested with this configuration in the near future.

Using the variable coherence length laser diode, the air paths were adjusted to give zero OPD (see Fig. 2 for distances) and low-current laser diode fringes were seen, then eventually broadband fringes from the xenon lamp were found with both a linear vertical polarizer and a narrowband (10 nm) yellow filter in place. The fringe intensity and contrast were weak (near the detection limit of the camera) due to the low coupling efficiency into Fiber B. Without the gel/cover slip patch on Fiber B, no detectable power could be seen by the camera. Therefore, the gel/cover slip is essential for low light levels.

PHASE LOCKING A MODULATED FIBEROPTIC INTERFEROMETER

As a separate experiment, a shorter pathlength (3 m each) Mach-Zehnder interferometer was built using bare single-mode polarization preserving Newport F-SPA fiber. This arrangement is shown in Fig. 3. The modulator is a Canadian Instrumentation and Research product with an oval fiber path, composed of two passive hemi-cylindrical ends with two planar piezoelectric sections in between. The fiber is wrapped seven turns around this path. It is epoxied over about 2/3 of the length of each piezo crystal and also, for keeping it in place, at the hemi-cylindrical ends. The modulator comes with a printed circuit driver board attached, with a sawtooth generator (adjustable frequency from 125-667 Hz), op amp front end for accepting a signal input, and a single transistor high voltage driver connected to the piezo crystals. A separate dc high voltage power supply is needed.

The total length of the fiber around the interferometer plus the two leads was measured indirectly by wrapping a thread along the same path as the fiber. This length was found to be 2.914 m after the ends were cleanly cleaved. The fiber of the other arm was therefore cleaved to be as close as possible to this length. This was achieved by stretching the fiber in a straight path (with moderate tension), cleaving one end, measuring the remaining length, then cleaving a calculated length from the other end. This gave a total length of 2.915 m, so the two branches were assumed to be within 1 mm of each other. Unfortunately, the last cleave was not clean, reducing coupled power, so later another 25 mm was cleaved from both fibers.

Modification and Characterization of Phase Modulator -

For the phase locking experiment later, the driver board needs to be able to accept two signal inputs, one for the dithering oscillator and one for the locking feedback signal from the lock-in amplifier. Therefore, the printed circuit board was modified to add another summing circuit to the inverting input of the front-end op amp. Figure 4 shows the original circuit with modification; the new input is labeled S2.

With the electronic modification completed, the fiber interferometer was assembled and clear output fringes were obtained with the green HeNe laser. This fringe pattern was expanded with a negative focal length lens, and the central fringe detected by a 2 x 3 mm silicon pin photodiode and transimpedance amplifier (10 Mohm feedback resistor).

Bare Fiber Interferometer for Phase Lock Experiments

6/21/92
DAC

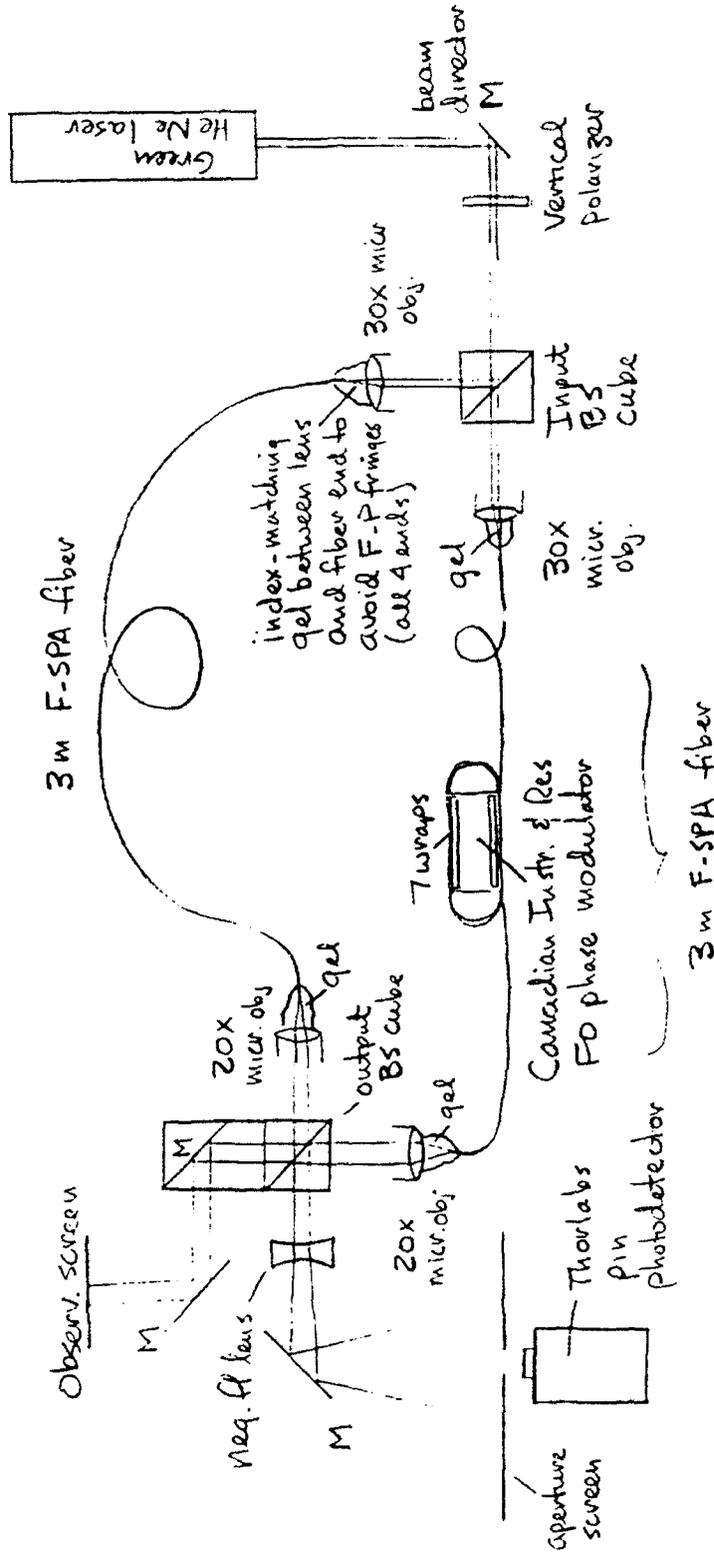
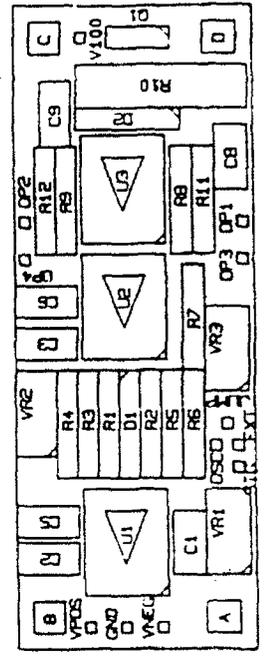
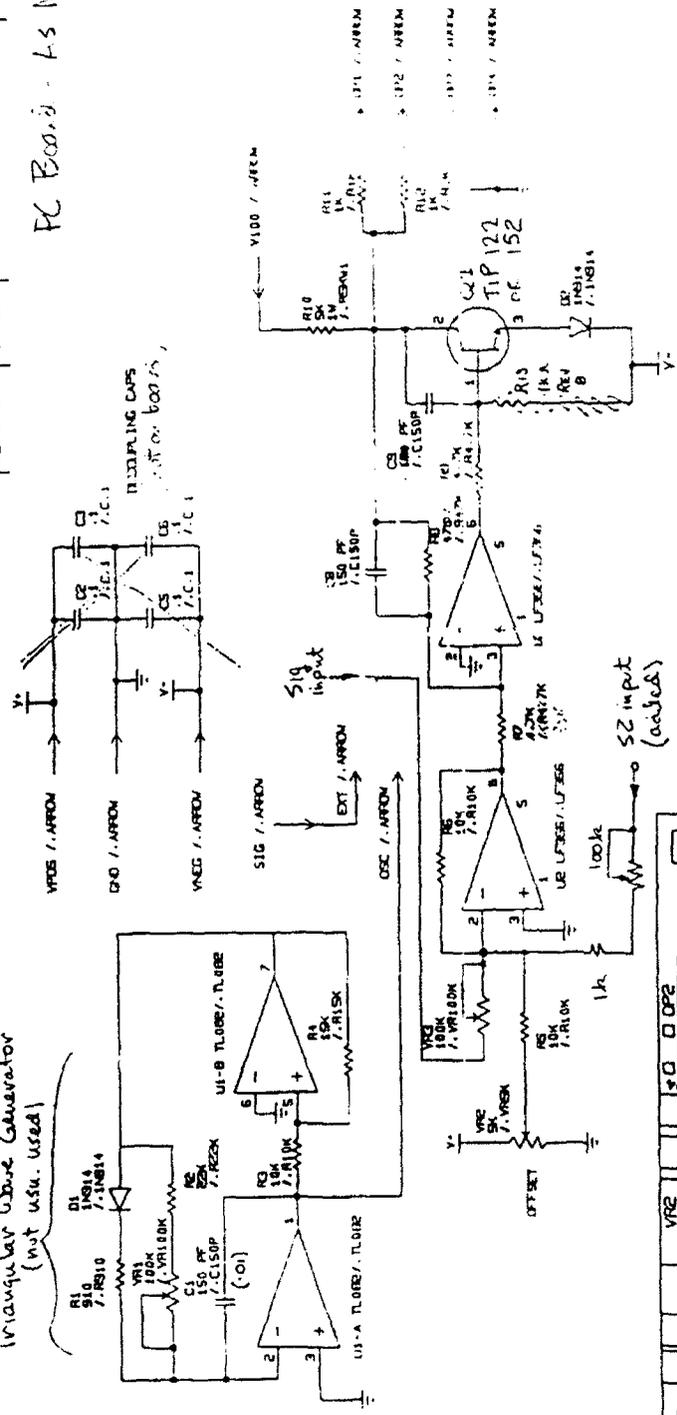


Fig. 3 - Separate interferometer for phase-locking studies and characterization of modulator.

DATE	SYM	REVISION RECORD	AUTH.	DRWN	CK.

PC Board - LS Modification
6/3/92
DLC

Triangular wave Generator
(not usu. used)



CIR LTD
OSC2 PCB SEPT/88
LAYER 5, SILKSCREEN

CANADIAN INSTRUMENTATION AND RESEARCH LTD		TIME	PIEZO OSCILLATOR SCHEMATIC	REV
MATL.		DATE	#2	8
FINISH		FLK		
		DRWN		
		DECKED		
		APPR		
		SIZE	B	
		SCALE		
		FORM. NO.	C502	
		REV	915EH-1	
		SHEET	OF	

Fig. 4 - The modification made to the modulator printed circuit board to allow two inputs.

One key parameter describing the performance of a phase modulator is the phase change in radians per volt of drive. This was measured by observing the fringe pattern while imposing a ramp voltage on the drive electronics to the piezoelectric crystal. The fringe changes are shown in Fig. 5 for a 40 V ramp in piezo voltage. From this plot, two features can be calculated:

1. There are 24 1/2 full fringe shifts over the 40 V, so the modulator sensitivity is:

$$24.5 \times 2\pi/40 = 3.85 \text{ rad/V.}$$

This is very acceptable sensitivity for seven wraps of fiber.

2. Since the full range of the drive voltage is 0-180 Vdc, this corresponds to a fringe range of:

$$24.5 \times 180/40 = 110.3 \text{ fringes.}$$

This range should be more than adequate to cover the fringe drifting that may be encountered over several minutes of operation.

Another key parameter of a modulator is its frequency response. This was measured by providing a constant input voltage (30 mV pk-pk) to the driver board op amp and observing the peak-to-peak fringe excursion while changing the frequency of the drive voltage. The dc offset voltage (approximately 50 V) was slightly changed on each observation to stay at the point of maximum slope, and thus maximize the fringe excursions. The peak-to-peak intensity change of the fringe excursion was approximately 0.25 of a complete minimum-to-maximum fringe shift.

The first measurement of the modulator's frequency response gave an unusually low 3 dB rolloff frequency of about 50 Hz. It was later discovered that this frequency was due to the limited frequency response of the photodetector (caused by a capacitor placed in the detection circuit to reduce noise), not the modulator. Thus, the experiment was redone in the author's lab at the University of Utah after returning from the summer assignment, again using the same procedure as described above except that a high speed photodetector was employed. The detector had a frequency response of greater than 1 MHz, determined by observing very short laser diode pulses with the detector; this value is much higher than the modulator frequency response, and therefore is not a limiting factor.

The results of the new modulator measurement are given in Fig. 6 as a log-log (Bode) plot. The 3 dB frequency is the frequency at which the response (modulation depth) is reduced to one-half

6/7/92
DAC

FRINGE SHIFT VS. PIEZO. DRIVE VOLTAGE
Canadian Instr & Res. Modulator; 7 fiber wraps F-SPA
3.85 $\text{m}\lambda/\text{V}$; 1.63V for 1 fringe shift

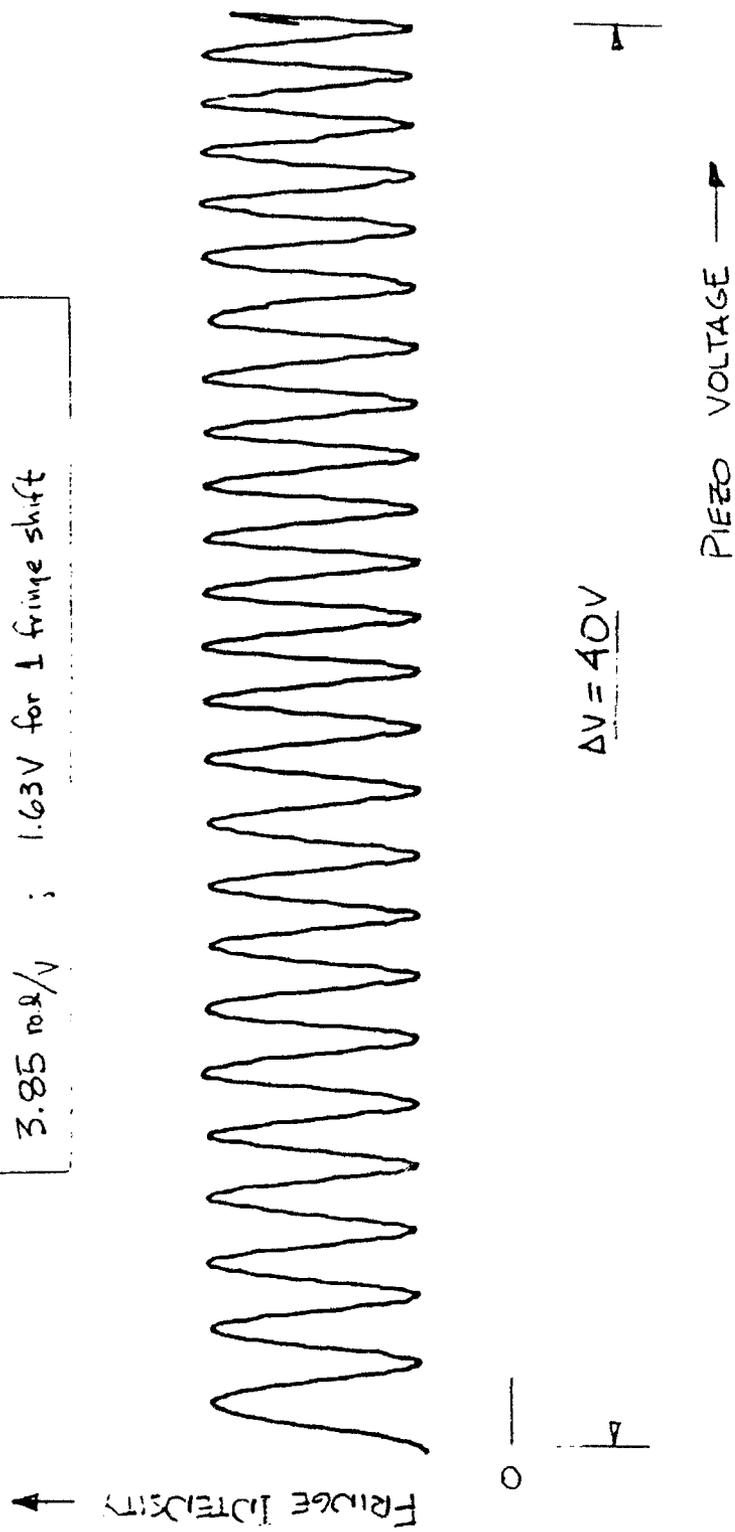


Fig. 5 - Fringe shifts for a 40 V ramp in drive voltage.

Frequency Response of Piezo Modulator

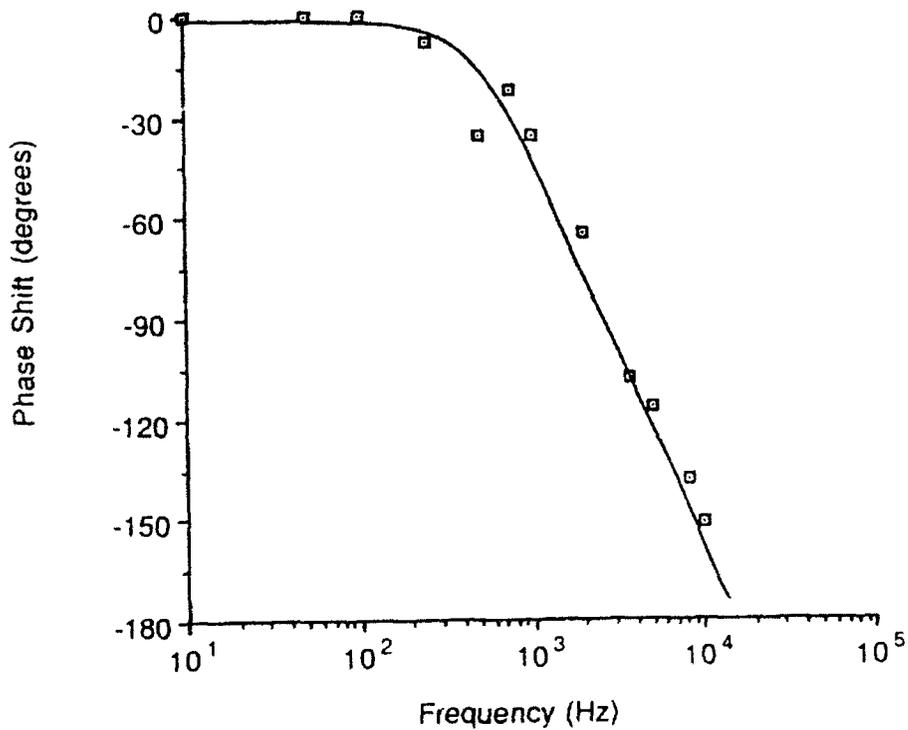
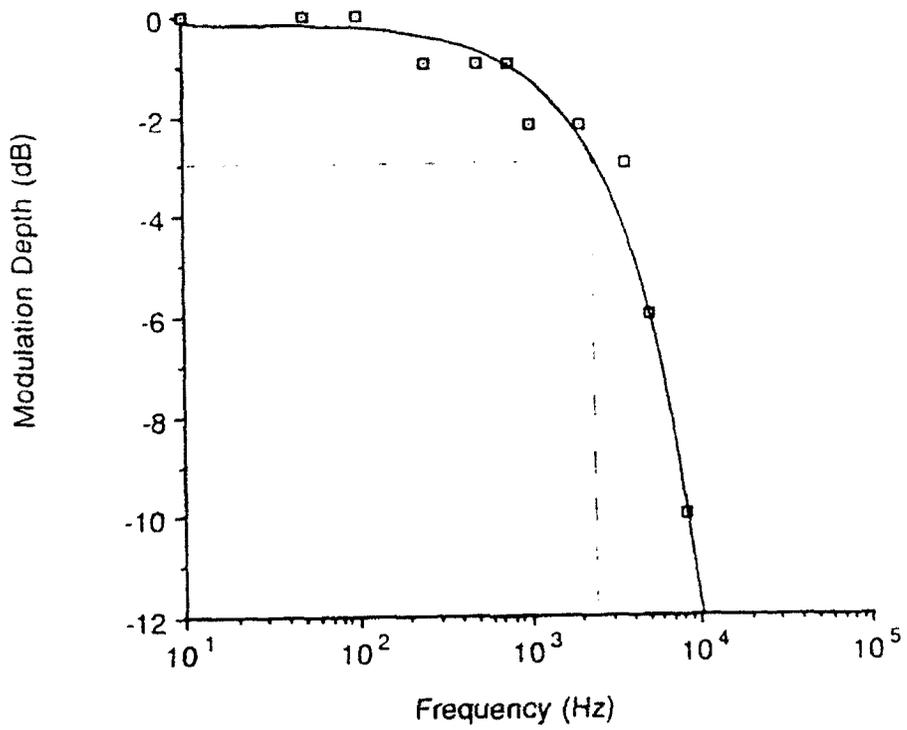


Fig. 6 - Frequency response curves for phase modulator. Input is to the driver board op amp.

of its low frequency value. On Fig. 6 this can be seen to be about 2.5 kHz. This frequency is consistent with the phase shift data; the 3 dB frequency is the frequency at which the phase lag is -90° with respect to the drive voltage, which can be seen to be about 2.5 kHz on the phase plot.

The modulator frequency response is therefore high enough to be operated in the phase lock experiments described next.

Phase-locking the Interferometer -

To achieve phase lockup, the electronics configuration shown in Fig. 7 was assembled. As seen in the figure, the key module was a Stanford Research SR510 lock-in amplifier, which provided both the dithering frequency (near 100 Hz) to the modulator, and the feedback control signal to lock the phase.

Two phase lockup experiments were completed. In the first, no special end preparations were used at the fiber ends. The graph of fringe drifts vs. time (over two 15 minute periods) for both the locked and unlocked cases is shown in Fig. 8.

For the unlocked case, the fringe drift due to ambient changes in the fibers is clearly seen. There are about 2 fringe shifts per minute. For the locked case (where the sign of the feedback signal gave a lock onto the fringe minimum), the fringe drift was much reduced, but some variation (estimated at 13%) is still visible. Over the 15 minute period, the dc level to the piezo changed by about 7 V, corresponding to a net 4.3 fringe drifts in one direction over this time.

The cause of the remaining variation under locked conditions was later discovered to be due to drifts of Fabry-Perot fringes caused by reflections from the ends of the fibers, and between the fiber ends and the laser output mirror. This was discovered by plotting the intensity output for one fiber only (the one wrapped around the modulator) as a function of drive voltage. This plot is shown in Fig. 9. Comparing these results to Fig. 4 shows that Fig. 9 has : a) smaller fringe contrast, and b) a fringe sensitivity exactly twice that of Fig. 4. Both of these findings are consistent with Fabry-Perot fringes in the fiber, due to small (4%) Fresnel reflections at the fiber ends and round-trip path multiple reflection interference.

To avoid the Fabry-Perot fringes, index matching gel (from Math Associates) was placed in the space between the microscope objective lens face and the fiber ends. This considerably reduced the Fresnel reflections. Then the phase lock experiment was repeated, with the results shown in

Phase Lock Electronics

6/21/92 DAC

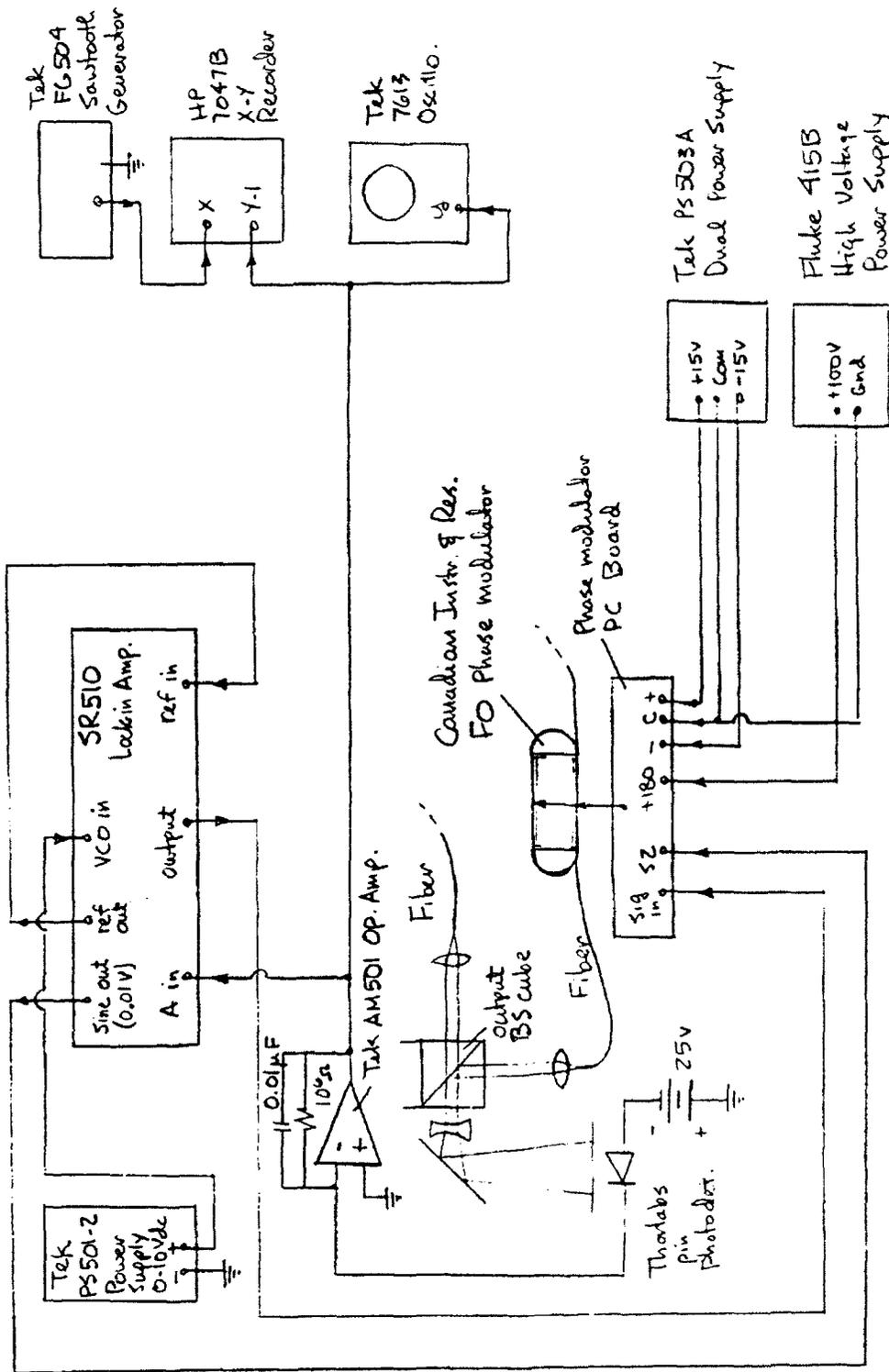


Fig. 7 - Electronic connections used to phase-lock the fiber interferometer.

6/7/92
DAC

FRINGE DRIFT FOR PHASE-LOCKED VS.
UNLOCKED FIBER INTERFEROMETER
3 m F-SPA fiber each branch; Candian Instr.
& Res. piezo. w/actuator; SR510 lock-in amp.

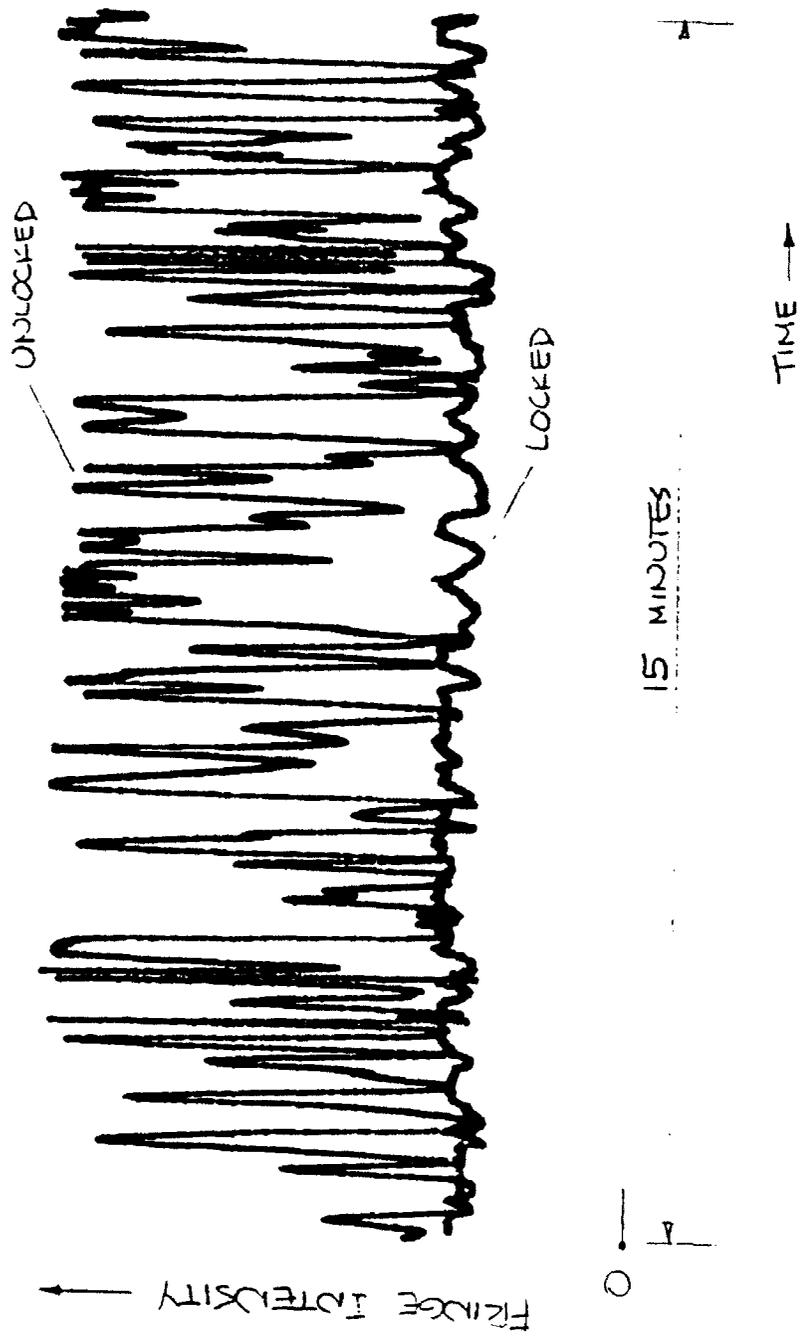


Fig. 8 - Fringe drift for locked and unlocked cases.

6/11/92
DAC

INTENSITY THROUGH ONE FIBER
(A) VS. PIEZO DRIVE VOLTAGE;
Carrier: I.E.L. Modulator; Turns: F-SPA
3.93 rad/V ; 0.80V for 1 F-P jump shift

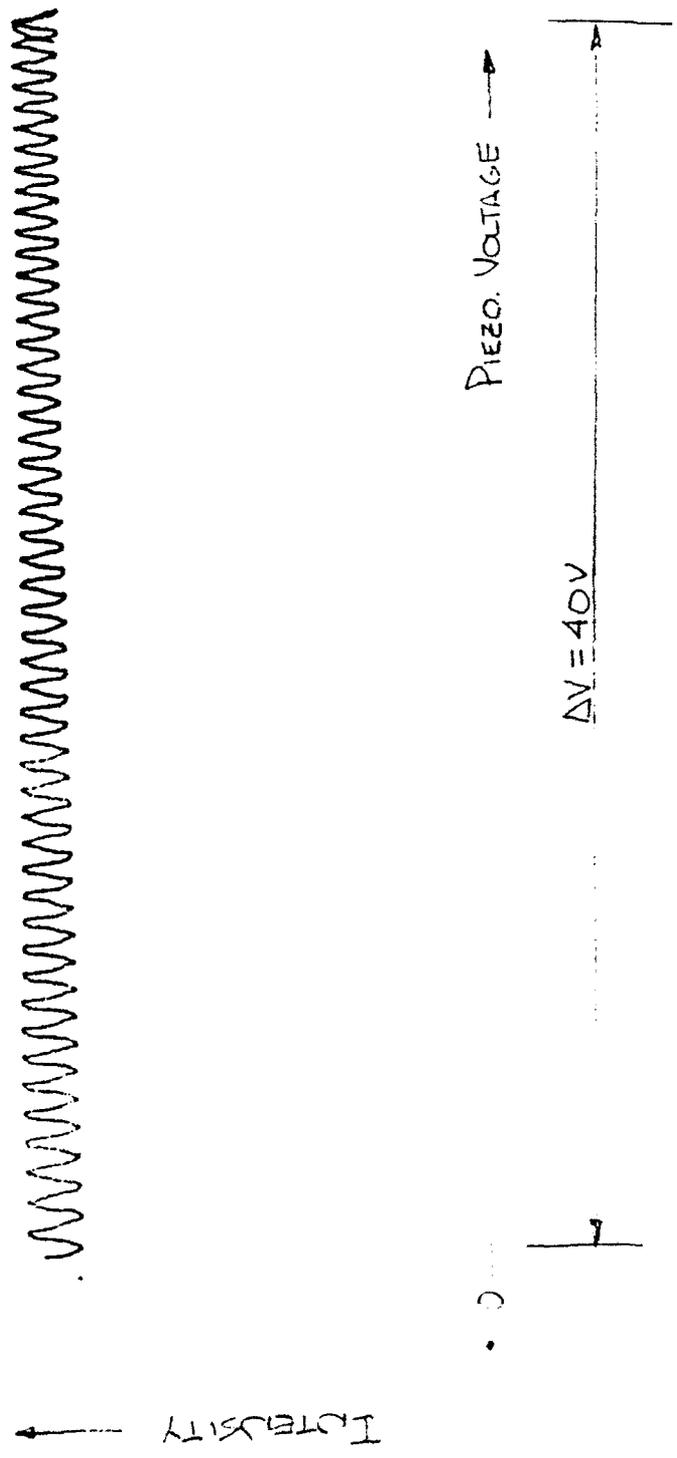


Fig. 9 - Fringe shift for one fiber only, for a 40 V ramp in drive voltage. The fringes are of a Fabry-Perot type.

6/20/92
DAC

FRINGE DRIFT FOR PHASE-LOCKED VS.
UNLOCKED FIBER INTERFEROMETER
3m F-SPA fiber each branch; ends gelled;
CIIR piezo modulator; SR510 lock-in Amp.

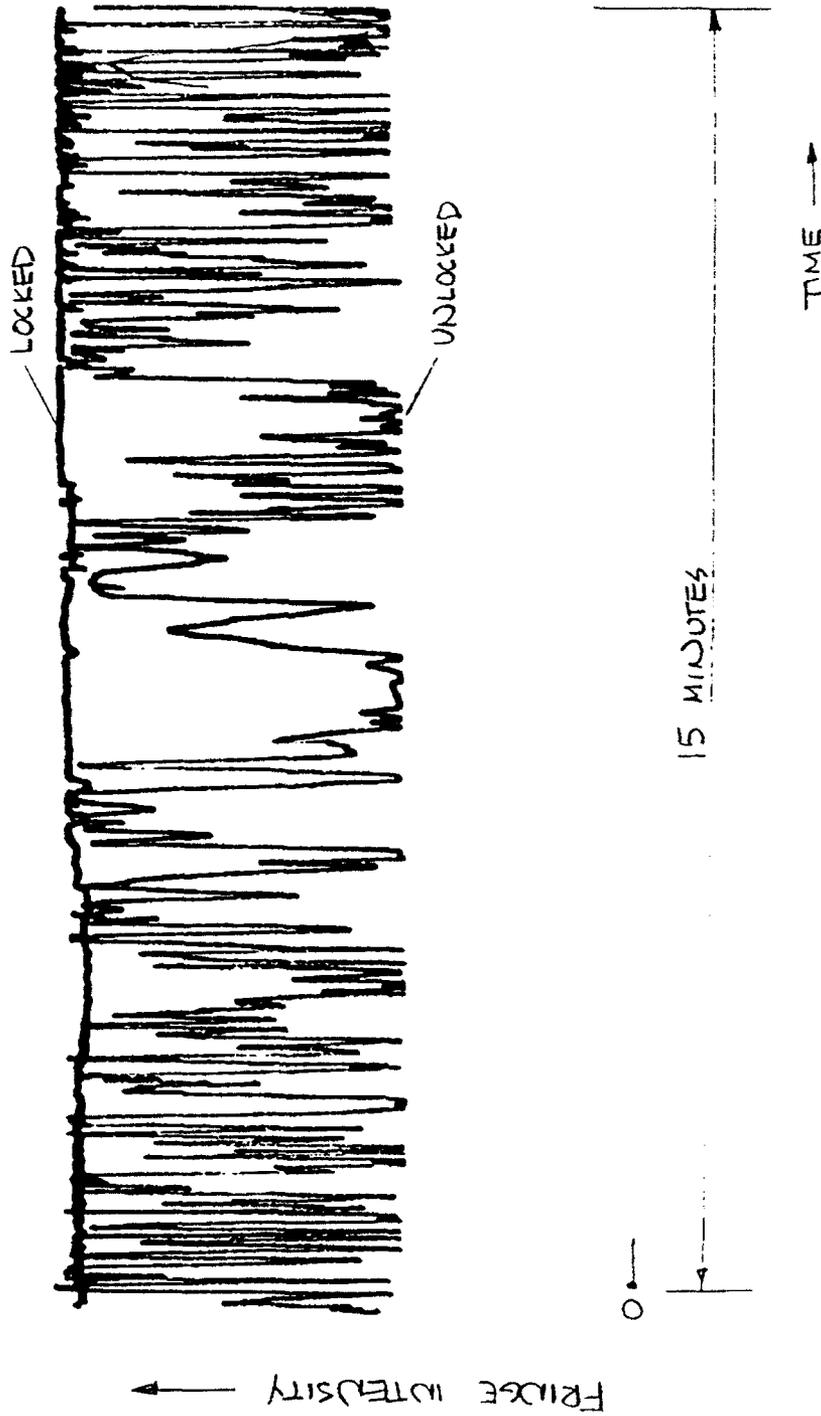


Fig. 10 - Fringe drift for locked and unlocked cases, with gel applied to fiber ends to avoid Fabry-Perot fringes.

Fig. 10. It is clear that the variation under the locked case is now much smaller, although there is still some evident, probably due to residual amplitude fluctuations in the laser power or coupling efficiencies into the fibers. The settings on the SR510 lock-in amplifier which proved to give the best phase lock were:

Frequency of modulation: 102.4 Hz
Phase: -83.9°
All filters (BP,f,2f): in
Sensitivity: 2 mV
Dyn Res: low
Expand: 1x
Rel: off
Offset: off
Time constants: pre = 1 sec; post = 1 sec.
Sine wave reference

CONCLUSIONS

Two tasks were achieved this summer:

- 1) Broadband interference fringes were found in an intermediate-step interferometric setup in the laboratory. This was a side-by-side arrangement with two subaperture telescopes observing the same quasi-point source and then passing the light to the two 60 m long fiber paths before recombination by beamsplitters.
- 2) A phase-lock condition was obtained in a test fiber interferometer using a piezoelectric fiber modulator and a lock-in amplifier system. This "internal metric" phase-locking condition is important for the eventual application of the interferometer for imaging purposes. The modulator was characterized in sensitivity and frequency response, and it proved to be well suited for use with this fiber interferometer.

ACKNOWLEDGMENTS

This work was supported by the AFOSR Summer Faculty Program. The help of Phillips Lab and Rockwell Power Systems personnel is gratefully appreciated, including Gary Loos, Andrew Klemas, Jim Rotge, and Bob McBroome.

ANALYSIS OF ONYNEX AND MSRP SEISMIC REFRACTION DATA IN NEW ENGLAND

**John Ebel
Associate Professor
Department of Geology and Geophysics**

**Andrew Paulson
Graduate Student
Department of Geology and Geophysics**

**Boston College
Chestnut Hill, MA 02167**

**Final Report for:
Summer Research Program
Phillips Laboratory**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.**

September 1992

ANALYSIS OF ONYNEX AND MSRP SEISMIC REFRACTION DATA IN NEW ENGLAND

John Ebel
Associate Professor
Department of Geology and Geophysics

Andrew Paulson
Graduate Student
Department of Geology and Geophysics

Abstract

Using interactive two-dimensional seismic raytracing techniques, seismic velocity models of the crust and upper mantle in New England were constructed from data collected jointly by Phillips Laboratory and Boston College from two seismic refraction experiments. One was from the 1984 Maine Seismic Refraction Experiment (MSRP), with the shotpoints in southcentral Maine and the receivers stretching from Rumford, Maine into the White Mountains in New Hampshire. The second was from the 1988 Ontario-New York-New England Seismic Refraction Experiment (ONYNEX) along a 200 km profile from western New York through Vermont and into southern New Hampshire. From the MSRP data the lower crust and Moho discontinuity in Maine were found to vary noticeably across the Norumbega Fault, with a significantly deeper Moho to the west of the fault. From the ONYNEX data the previously reported ramp feature, separating the Grenville basement in the Adirondack Mountains and the Paleozoic basement in the Northern Appalachians, was found. The ramp is inferred to dip from the surface near the Vermont-New York border to 17 km depth beneath the Vermont-New Hampshire border. The configuration of the ramp indicates that it controlled the emplacement of the geology above it during Paleozoic and Mesozoic time. The results of this study suggest models for ancient continental zones which may be applicable to those in other parts of the world.

INTRODUCTION

In 1984 and again in 1988 the Air Force Geophysics Laboratory (now Phillips Laboratory) and Boston College (called here PL-BC) jointly participated in carrying out piggyback experiments as part of large-scale seismic refraction studies in New England. The two experiments, the Maine Seismic Refraction Experiment (MSRP) in 1984 and the Ontario-New York-New England Seismic Refraction Experiment (ONYNEX) in 1988, were carried out by the U.S. Geological Survey (USGS) to determine the seismic velocity structure of the crust across an ancient continental collision zone. Other participants in parts of MSRP and ONYNEX were the Massachusetts Institute of Technology (MIT) and the Geological Survey of Canada. The PL-BC participation in these experiments was designed to expand the areal coverages of these experiments, to test field methods for recording seismic refraction data, and to provide a detailed data set for waveforms studies of explosions in areas of complicated geologic histories. The ultimate goal of the PL-BC research was to improve the methodologies for monitoring nuclear testing treaties in remote parts of the world by learning to model regional seismic waveform data from a well-studied area.

The area of the surveys and the locations of stations of the PL-BC piggyback experiments are shown in Figure 1. For MSRP the PL-BC stations for the first night of the experiment were located in the White Mountains of Maine and New Hampshire to the west of the shots which were fired that night. It is these data which we analyzed during the summer, 1992 research effort. For ONYNEX data primarily from the second night of the experiment, a line of receivers from just northwest of Westport, New York to Manchester, New Hampshire, were modeled. Seismic raytracing through models of the seismic velocity structure of the crust along each of the profiles was used to match the arrival times of the major body wave phases in the observed seismograms.

A number of crustal models for New England, based on the USGS data, have already been published for the MSRP and ONYNEX experiments. For MSRP early work based on first analyses of the data are found in Luetgert et al. (1987) and

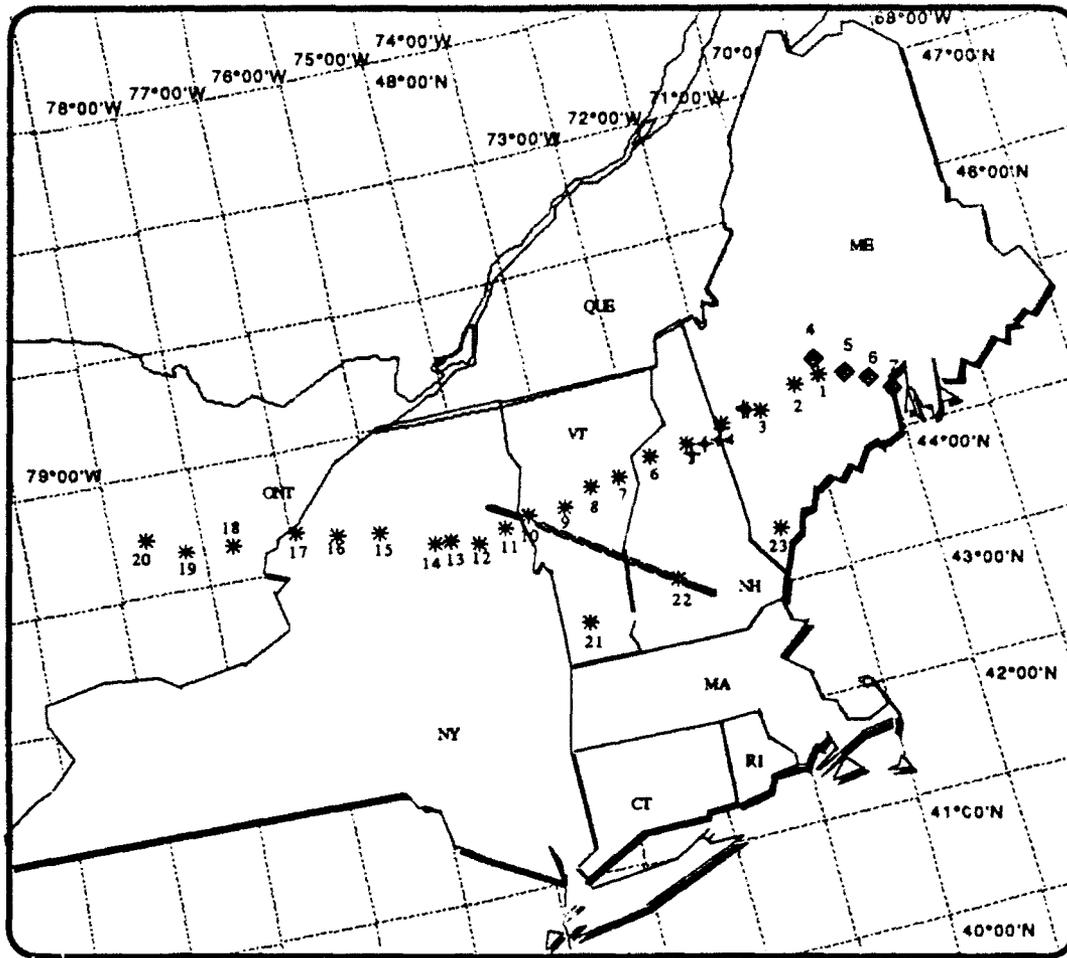


Figure 1. Map of the study area showing the locations of the sources and receivers for this analysis. The shotpoints and receivers from the 1984 Maine Seismic Refraction Experiment used in this study are indicated by the diamonds and crosses respectively. All of the shotpoints for the 1988 Ontario-New York-New England Seismic Refraction Experiment are shown (the stars) as is the trend of the seismic line analyzed in this study.

Klemperer and Luetgert (1987). More recently, a detailed model of the crust for the along-strike profile in central Maine was published by Hennen et al. (1991). The data from the long, cross-strike profile from Canada to coastal Maine are still being analyzed by the USGS during the summer of 1992 (J. Luetgert, personal communication, summer, 1992). For ONYNEX a detailed model for the eastern half of the experiment, stretching from Maine into central New York state, was published by Hughes and Luetgert (1991). The results reported in these studies are important because they provide constraints and starting models for the analysis of our data sets. These studies are also important because we chose to use with our observations the same analysis technique (ray tracing using the computer programs developed by Luetgert, 1988) as was used in those previous studies. About halfway through the summer, J. Luetgert provided an improved version of his code, in this case ported to Macintosh computers (Luetgert, 1992). The use of this code greatly facilitated the modeling of the very complex structure sampled by the PL-BC data from the ONYNEX experiment.

GEOLOGIC SETTING AND TECTONIC HISTORY

Because the present seismic structure of New England is a reflection of its past tectonic history, we present a brief summary of the geologic development of New England relevant to the areas studied here. This summary is abstracted from a number of sources, most notably Taylor and Toksoz (1979) and Press and Siever (1982). A summary map of the geology together with the locations of the two study areas are indicated in Figure 2.

During the late precambrian the eastern edge of North America rifted from a landmass to the east and evolved into a passive margin. The edge of the continent in New England at that time ran from what is today western Massachusetts north through central Vermont and then northwest across the northern tip of New Hampshire and along the northern third of the state of Maine. After a long period of drifting a series of collisional events started in New England with the Taconic orogeny about 480 m.y. ago. During this orogenic episode, oceanic sediments from the continental shelf and continental slope were thrust westward and northward upon the eastern edge of North America, forming

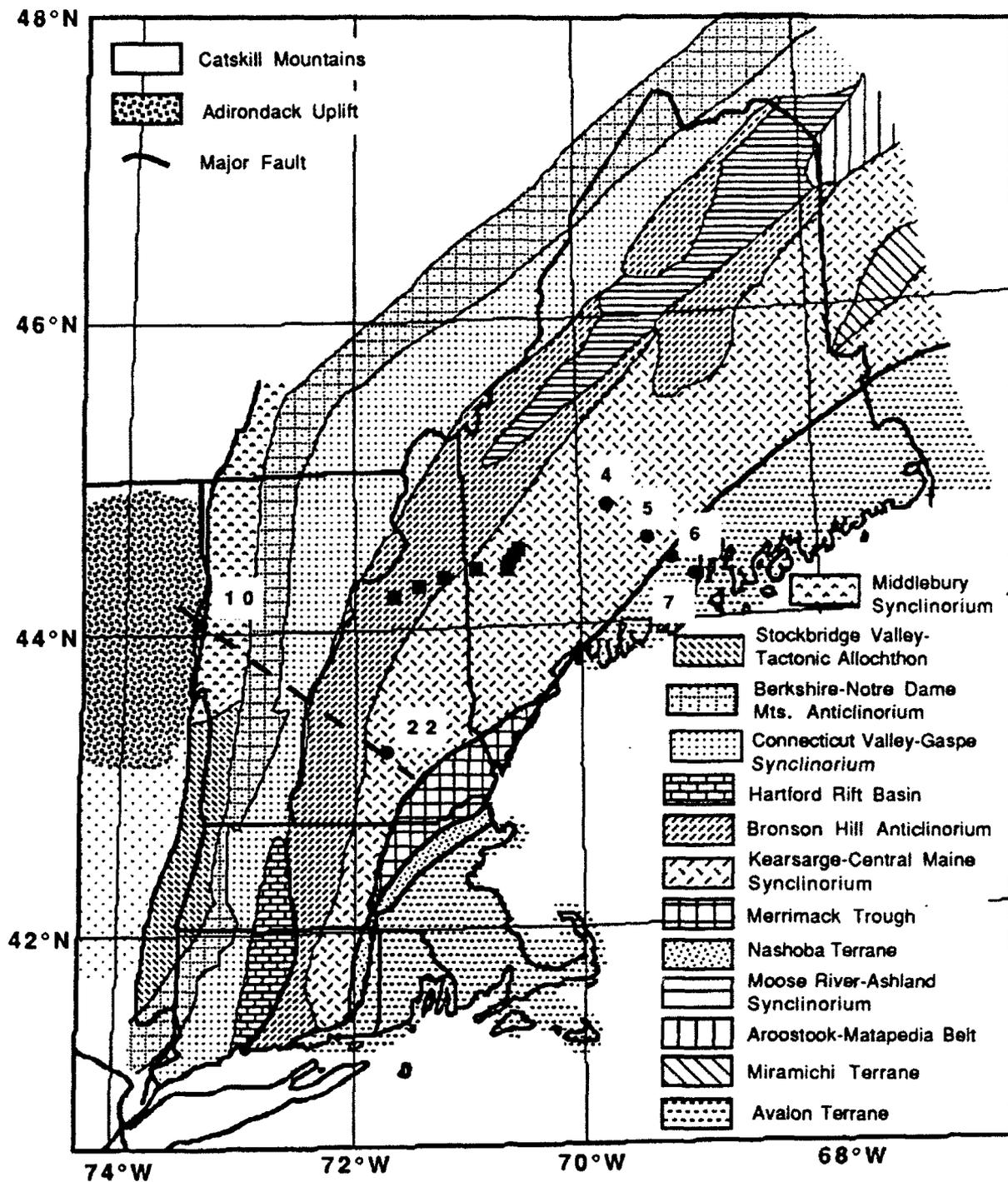


Figure 2. Map of the major geologic terranes and faults in New England. Shotpoints 4, 5, 6 and 7 from the MSRP and the receivers which recorded those shots (squares) are indicated as are shotpoints 10 and 22 from the ONYNEX along with the trend of the PL-BC profile (dashed line).

the Taconic Mountains among other structures. A part of the continental edge also broke off and was thrust up and west, exposed today in various bodies such as the Green Mountains in Vermont. Following the Taconic orogeny there occurred a series of collisions between North America and several island arcs and continental fragments from the east, along with the emplacement of a number of plutons of various ages. In New England, the Norumbega Fault in Maine and the Clinton-Newbury Fault in Massachusetts with its extensions in Connecticut are mapped as the surficial sutures between Avalonia, one of the largest of these continental fragments, and North America. The Acadian orogeny, which took place about 380 m.y. ago, was a brief but intense collisional orogeny, probably of an island arc with North America. The Alleghenian orogeny, which occurred about 300 m.y. ago, was probably associated with the final suture of the African land mass to the eastern North American margin. Subsequent to the Alleghenian orogeny Africa was attached to North America during the period of the existence of the supercontinent of Pangea. The breakup of Pangea in New England began during the Triassic period (about 210 m.y. ago) with the opening (which later failed) of a number of continental rift basins. During the Jurassic (about 180 m.y. ago) new set of basins formed to the east, and these successfully opened to form the Atlantic Ocean. The late Mesozoic (about 120-100 m.y. ago) in New England saw the emplacement of the major plutons which make up the White Mountains. These are thought to have been due to the passage of North America over a mantle hot spot. Finally, during the Cenozoic (since about 80 m.y. ago) the eastern margin of North America has once again evolved into a passive margin, a state which still exists today.

DATA SETS FOR THIS STUDY

For both the MSRP and ONYNEX experiments the source shots were located and detonated by the USGS. The shotpoints were separated by about 30 km along the refraction lines set up by the USGS, with a subset of all shot points being detonated during each night of shooting. In both the MSRP and ONYNEX experiments the shots ranged in size from 1600 lbs. to 4000 lbs. of ammonium nitrate with one additional 6000 lb. shot in Maine during ONYNEX. The shots were detonated in 8-inch diameter holes, drilled 49-55 meters in the bedrock to ensure

good coupling. One shot point each during MSRP and ONYNEX was in a water-filled quarry. Data from these quarry shots were not used in this study.

The instrumentation and recording configurations for the PL-BC experiments were different during the MSRP and ONYNEX experiments. For the first night of shooting during the MSRP (the data set analyzed here), five PL Terra Technology DCS-302 recorders with 3 Hall-Sears HS-10-1B 1-Hz geophones were spaced about 15 km apart from Rumford, Maine westward to Crawford, New Hampshire. Just west of Rumford Boston College installed at about 1 km station spacing 5 Sprengnether MEQ-800 analog recorders with vertical HS-10 1-Hz sensors and 2 digital Sprengnether DR-200 recorders with 3 Teledyne S-13 1-Hz sensors. For the ONYNEX data set analyzed here PL-BC along with MIT installed a line of seismographs at about 5 km station spacing from Manchester, New Hampshire to just northwest of Port Henry, New York. The primary purpose of this line was to record two shotpoints from near either end of the line. PL put out 25 Terra Technology recorders with Sprengnether S-6000 triaxial geophones, while the Boston College-MIT data were recorded on 9 seismographs, 5 of which were MEQ-800 analog seismographs which were connected to HS-10 seismometers. The other seismographs were DR-200 digital seismographs connected to S-6000 seismometers. All of the instruments during both experiments were calibrated either to WWV radio or GOES satellite time.

As the first step in the analysis of the data, the digital waveforms for both experiments were assembled on the PL VAX computer. The waveforms were then read into the computer program ICON, developed by Dr. John Cipar at PL, where the first arrival times and the arrival times of other important P wave phases were read. Using ICON we constructed record sections of the data, facilitating identification of phases from one waveform to the next. An example of a record section from the ONYNEX data set is shown in Figure 3. We then added to the set of arrival times from the digital data the times of arrival of first and later P-wave arrivals from the analog records. These were read with a visual ocular graded to a 0.1 mm scale. The reading accuracy of the digital waveforms was 0.01 sec, while for the analog waveforms it was 0.02 to 0.04 sec.

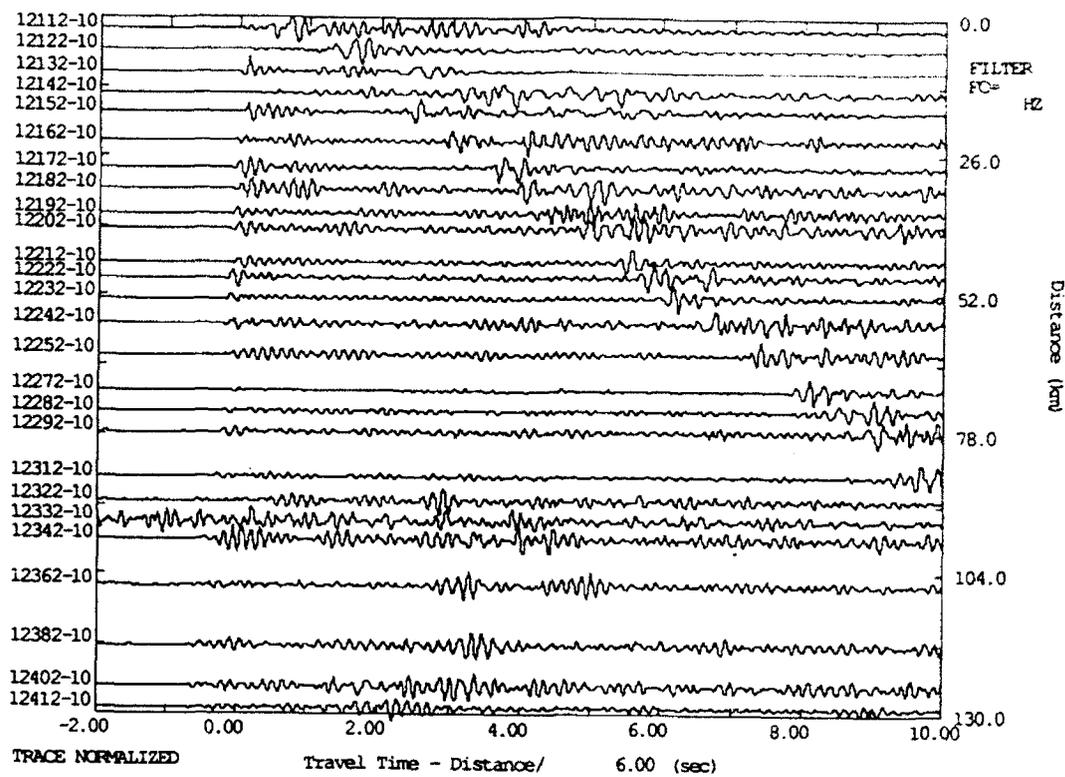


Figure 3. Example of one of the PL-BC seismic sections from the ONYNEX experiment. The traces are from shotpoint 10 to the southwest, with the 0 distance being at shotpoint 10. The time axis represents travel time at a reduced velocity of 6.0 km/s. The amplitudes of the traces are individually normalized.

MODELING RESULTS

MSRP Data Set

The set of seismograms from the MSRP experiment contained a number of P-wave arrivals which were used in the ray-tracing analysis. In many cases the first arrivals were rather emergent, indicating that these arrivals were head waves or waves guided along the top of faster velocity layers in the upper crust. Several later arrivals in the P waveforms were also analyzed. These generally fit the expected travel times for mid-crustal or Moho reflections. Thus, the arrival time data set could be used not only to measure seismic velocities in the upper crust but also to estimate the thickness of the crust in the study area.

The approach in fitting the observed travel times was first to construct a starting crustal model based on the results of Hughes and Luetgert (1991) and Hennett et al. (1991). Rays were then traced through this model from the appropriate source positions to each receiver location, and the predicted travel times were compared to the observed travel times. The crustal model was then changed and the ray tracing redone. This process was repeated a number of times by trial and error until a velocity model which closely matches the travel time data was found. The data were analyzed starting with that from the closest shotpoint and working progressively toward the farthest shotpoint. For the observations from each shotpoint, the first arrivals were matched first followed by the later arrivals in the data set. The deepest reflections, those from the Moho, were fit last for each shotpoint.

Plots of the crustal model which best fit the arrival time observations are shown in Figure 4. Several interesting features can be observed in this model. First, while a low velocity zone from the starting model is preserved in the upper crust, the data set analyzed here is too sparse to either confirm or deny the occurrence of such a feature. Second, the top of the lowermost crustal layer is quite deep between shotpoints 4 and 5 and the receiving stations. It is as deep as 29 km here, whereas it is between 21 km and 27 km deep throughout other parts of New England. The Moho depth of 37-40 km under the central part of this profile is

**Western Maine Crustal Models
PL-BC Data 9/25/84**

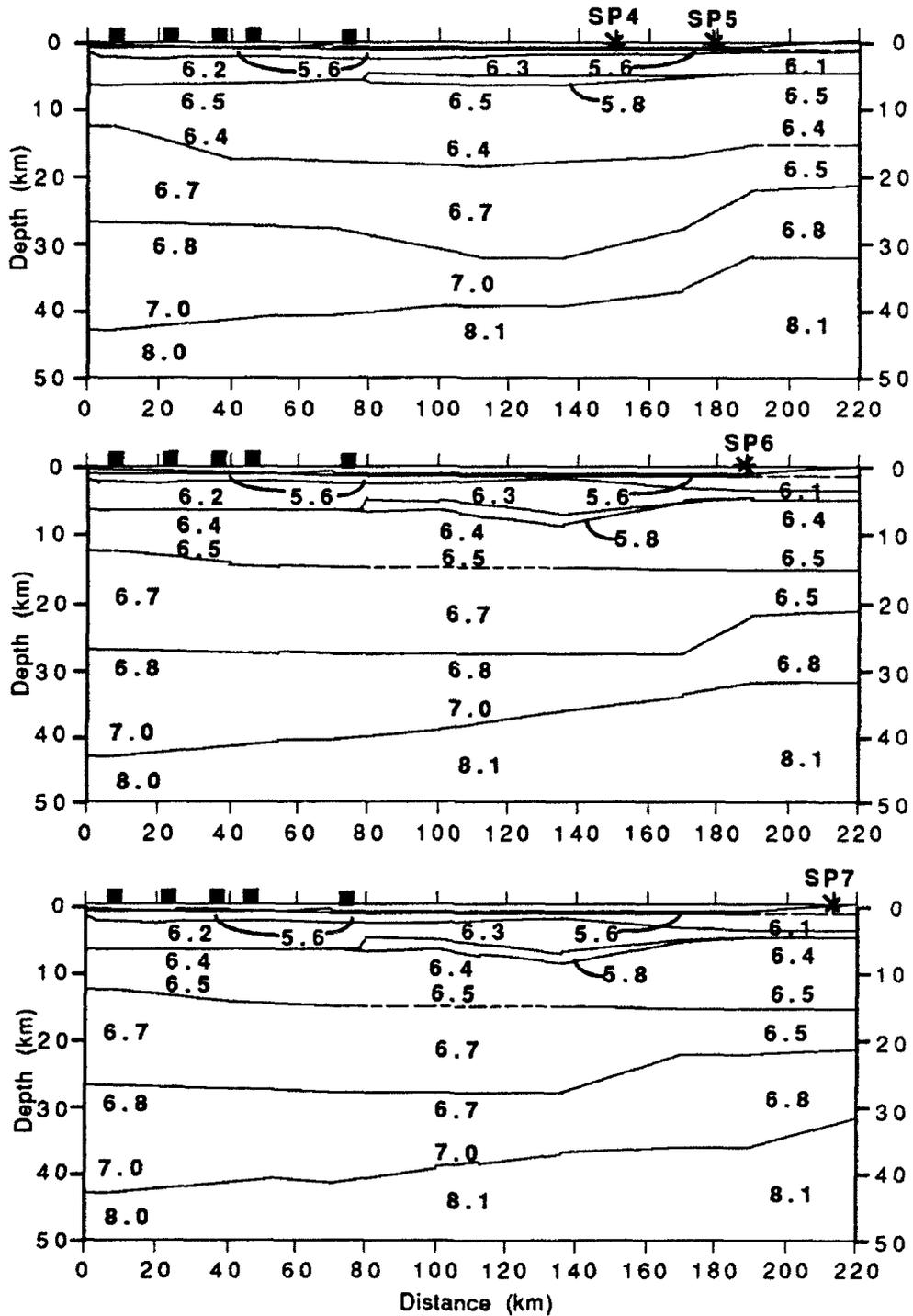


Figure 4. Crustal models for western Maine from the analysis of the MSRP data. Shown are the P-wave velocities in km/s at different points in the models. The shotpoint locations (stars) and receiver locations (solid squares) relative to the model are indicated at the top of each cross section.

quite consistent with the results of Luetgert et al. (1987) and Kafka and Ebel (1988) who reported a local depression in the Moho in this vicinity.

For shotpoints 6 and 7 there are some notable changes in the crustal model needed to fit the data. Specifically, the Moho is shallowed to a depth of 33-35 km near the shotpoints, and the top of the lowermost crust is raised to about 22 km. These changes occur to the east of the Norumbega Fault, and they suggest that the Norumbega Fault represents a major discontinuity in the structure of the lower crust. Curiously, such a discontinuity was not reported either by Klemperer and Luetgert (1987) or Stewart et al. (1987). However, the observations in this study sample a somewhat different part of the crust of Maine than those studies. Thus, it appears that strong lateral changes in the seismic velocity structure of the lower crust of Maine can occur over distances of a few tens of kilometers.

ONYNEX Data Set

Applying to the PL-BC ONYNEX data set the same analysis methods as were described above for the MSRP observations, we first found constraints on the layer thicknesses and velocities under southern New Hampshire and central Vermont and under the easternmost part of the Adirondack Mountains. These constraints were used in the creation of two seismic velocity models for the seismic line which we studied. One model was for the arrivals northwest from shotpoint 22 in Southern New Hampshire (an Appalachian crustal model), while the second was for the arrivals around shotpoint 10 in the Adirondacks of New York (an Adirondack crustal model). Major differences in the seismic velocity structure between these two regions had already been reported by Hughes and Luetgert (1991). A single crustal model was constructed from these two distinct models by ramping the Adirondacks underneath the Appalachians, as previously recognized from the 1980 COCORP profile (Brown et al., 1983) and from the Hughes and Luetgert (1991) study. The seismic velocities and positions of the layer boundaries in the constructed model were varied until rays traced through the model best fit the observed arrival times. The interactive two-dimensional seismic raytracing program MACRAY (Luetgert, 1992) greatly facilitated the ray tracing processing. This analysis led to the development of a seismic velocity model that

is similar to that found by Hughes and Luetgert (1991) for the region just north of our study area.

The best fitting seismic velocity structure (Figure 5) is well constrained in the upper 20 km as the first arrivals were strong and relatively clear of noise and the source-receiver distances provided observations which primarily sampled the upper half of the crust. An insufficient number of strong/clear second arrivals representing reflections from the deeper parts of the crust and from the top of the mantle lead to more ambiguous readings of their arrival times, and therefore, the velocity structure below 20 km is not as constrained as that of the upper crust.

The northern end of the profile extends into the Adirondack mountains. The Adirondacks velocity crustal model consists of a low velocity layer of seismic velocity 5.45-5.50 km/s. This 0.5 km thick layer represents the weathered rock and ground cover. From 0.5 km to 18 km depth the Adirondacks are characterized by a high P-wave velocity of 6.30-6.55 km/s. This high P-wave velocity reflects the anorthositic composition of the Adirondacks, which consist of 1.0+ b.y. old Grenville basement rocks. The high crustal velocity rocks of the Adirondacks extend from the near surface, at the northwestern part of the profile, to a depth of 17 km under the Vermont-New Hampshire border. The Appalachian crust is also topped by a 0.5 km weathering layer, and below that the P-wave velocities increase from 6.03 km/s to about 6.3 km/s at 15 km depth. An upwarp in the top of the 6.65 km/s layer near the base of the ramp was needed to fit the first arrival time data from shotpoint 22 to the stations at the northwestern part of the profile. A local upward bulge in the Moho surface was also inferred from the arrival times of what were interpreted to be PmP reflected phases in the sections.

The major result of the analysis of the PL-BC ONYNEX data is the position of the ramp under the profile. The ramp is found to dip steadily from the Vermont-New York border downward to the Vermont-New Hampshire border where it flattens out considerably. The bottom of the ramp under this profile occurs approximately due south of the bottom of the ramp under northcentral Vermont found by Hughes and Luetgert (1991). Thus, the edge of the ramp beneath Vermont and New Hampshire strikes approximately due north-south, which is the

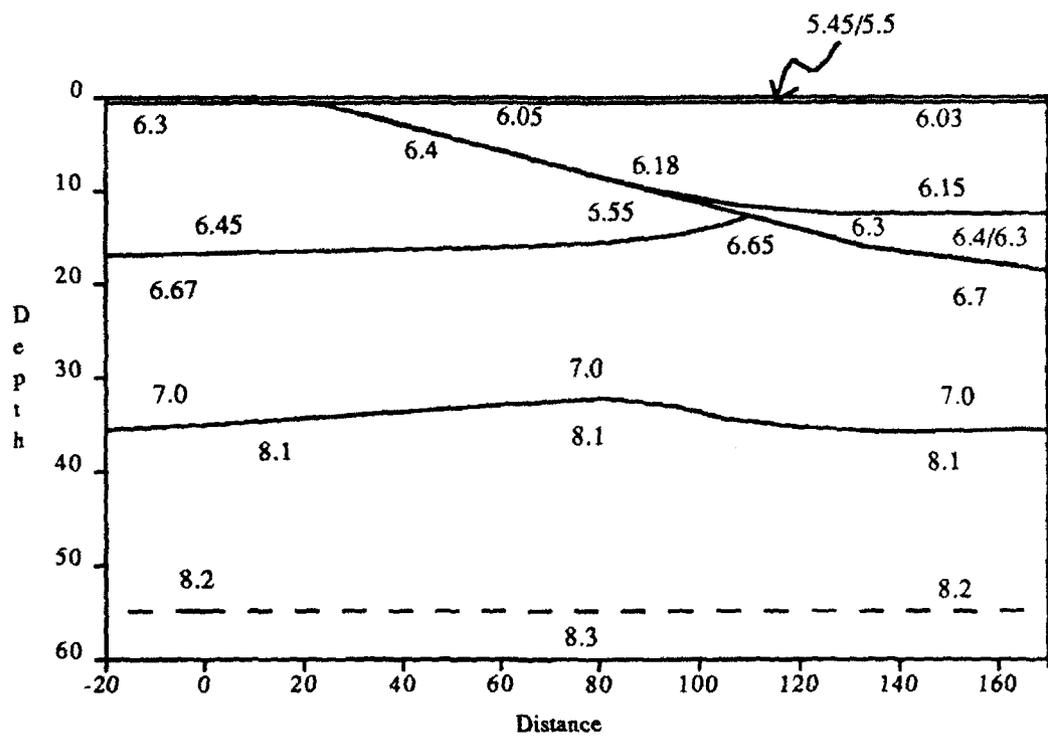


Figure 5. Crustal model for the PL-BC line from the eastern edge of the Adirondack Mountains to southcentral New Hampshire from the ONYNEX data. Shown are the P-wave velocities in km/s at different points in the model.

prevailing trend of most of the surface geology in this area. Given the ancient age of the rocks in the Adirondack Mountains compared to that of the Appalachian rocks, it appears likely that the shape of the Adirondack/Grenville massif has controlled the orientation of the emplacement of the geology and tectonic terranes in Vermont since late precambrian time.

IMPLICATIONS FOR NUCLEAR TEST BAN TREATY MONITORING

One of the most important aspects of using seismology to monitor nuclear test ban treaties, particularly using seismic stations located at regional distances (less than 100 km to 2000 km or so from the sources), is to understand the seismic wave propagation between the sources and the receivers (Pomeroy et al., 1982). Seismological studies in recent years have revealed strong complexity in the structure of the earth's crust in many places (Braile, 1991), and the variation of this complexity from place to place makes it difficult to find general rules to guess the seismic wave propagation characteristics for any particular source-receiver path. The results of this study demonstrate that, for an ancient continental collision belt, major crustal variations can occur over distances of tens of kilometers. In particular, the topography of the Moho can vary by several kilometers, and the interface between older continental shield rocks and younger, overthrust collisional terranes is a smoothly dipping structure with a geometry which is reflected in the overlying surface geology. These results give important clues about what crustal seismic structures may be found in other parts of the world, such as under the Ural Mountains and the Caucasus Mountains in Eurasia. They can be used to postulate possible crustal models for use in nuclear test ban treaty monitoring studies in these and other areas of the world.

REFERENCES

Braile, L.W. (1991). Seismic studies of the earth's crust, *Reviews of Geophysics, Supplement, U.S. National Report to International Union of Geodesy and Geophysics, 1987-1990*, 680-687.

Brown, L., C. Ando, S. Klemperer, J. Oliver, S. Kaufman, B. Czuchra, T. Walsh, Y. Isachsen (1983). Adirondack-Appalachian crustal structure: The COCORP northeast traverse, *Geol. Soc. Am. Bull.*, **94**, 1173-1184.

Hennet, C.G., J.H. Luetgert and R.A. Phinney (1991). The crustal structure in central Maine from coherency processed refraction data, *J. Geophys. Res.*, **96**, 12,023-12,037.

Hughes, S., and J.H. Luetgert (1991). Crustal structure of the western New England Appalachians and the Adirondack Mountains, *J. Geophys. Res.*, **96**, 16,471-16,494.

Hughes, S., and J. Luetgert (1992). Crustal structure of the southeastern Grenville province, northern New York state and eastern Ontario, *J. Geophys. Res.*, **97**, in press.

Kafka, A.L., and J.E. Ebel (1988). Seismic structure of the earth's crust underlying the state of Maine, in *Studies in Maine Geology*, Volume 1, Maine Geological Survey, 137-156.

Klemperer, S.L., and J.H. Luetgert (1987). A comparison of reflection and refraction processing and interpretation methods applied to conventional refraction data from coastal Maine, *Bull. Seism. Soc. Am.*, **77**, 614-630.

Luetgert, J.H. (1988). User's manual for RAY84/R83PLT interactive two-dimensional raytracing/synthetic seismogram package, *U.S. Geological Survey Open-File Report 88-238*, 52 pp.

Luetgert, J.H. (1992). MACRAY interactive two-dimensional seismic raytracing for the Macintosh, *U.S. Geological Survey Open-File Report 92-356*, 43 pp.

Luetgert, J. and C.E. Mann (1990). Avalon terrane in eastern coastal Maine: Seismic refraction-wide-angle reflection data, *Geology*, **18**, 878-881.

Luetgert, J., C.E. Mann, and S.L. Klemperer (1987). Wide-angle deep crustal reflections in the northern Appalachians, *Geophys. J.R. astr. Soc.*, **89**, 183-188.

Pomeroy, P.W., W.J. Best and T.V. McEvilly (1982). Test Ban Treaty Verification with Regional Data--A Review, *Bull. Seism. Soc. Am.*, **72**, S89-S130.

Press, F. and R. Siever (1982). *EARTH*, W.H. Freeman and Co., New York, 613 pp.

Stewart, D.B., J.D. Unger, J.D. Phillips, R. Goldsmith, W.H. Poole, C.P. Spencer, A.G. Green, M.C. Loiselle, and P. St. Julien (1986). The Quebec-western Maine seismic reflection profile: Setting and first year results, in *Reflection Seismology: The Continental Crust*, Geodynamics Series Volume 14, 189-199.

Taylor, S., and M.N. Toksoz (1979). Three-dimensional crust and upper mantle structure of the northeastern United States, *J. Geophys. Res.*, **84**, 7627-7644.

PREPARATIONS FOR NEUTRON SCATTERING INVESTIGATIONS
OF LIQUID-CRYSTAL POLYMERS

David M. Elliott, Ph.D.
Associate Professor
Department of Engineering

Arkansas Tech University
Russellville, AR 72801

Final Report for:
Summer Research Program
Phillips Laboratory
Edwards Air Force Base

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

September, 1992

PREPARATIONS FOR NEUTRON SCATTERING INVESTIGATIONS
OF LIQUID-CRYSTAL POLYMERS

David M. Elliott, Ph.D.
Associate Professor
Department of Engineering
Arkansas Tech University

Acknowledgement

The author would like to specifically acknowledge the efforts of Dr. Kevin Chaffee and Dr. John Rusek for their constant help and encouragement during this Summer Research Project. The technical competence of these Phillips Laboratory researchers is truly world class. It was a distinct pleasure to work with both of them on this extremely interesting and challenging research project.

PREPARATIONS FOR NEUTRON SCATTERING INVESTIGATIONS
OF LIQUID-CRYSTAL POLYMERS

David M. Elliott, Ph.D.
Associate Professor
Department of Engineering
Arkansas Tech University

Abstract

Researchers at the Phillips Laboratory at Edwards Air Force Base have been instrumental in the development and characterization of thermotropic liquid crystal polymers (LCPs). One of the most interesting and potentially valuable areas of research on these materials is the question of physico-chemical annealing. The mechanical and chemical properties of some of these materials are known to dramatically improved upon annealing. In order to realize the full potential of annealed LCPs, a firm understanding of the molecular structure, morphology, texture, and the dynamics and mechanism of formation of the annealed state must be obtained. Neutron scattering techniques are thought to be capable of providing much of the basic structural information necessary to develop that understanding.

During the summer of 1992, the author was privileged to have participated in the Phillips Laboratory LCP Research Program at Edwards Air Force Base. His work was primarily involved with preparations for a series of planned neutron scattering experiments. A preliminary search of the literature available about neutron scattering in general and small angle neutron scattering in particular was completed. The information obtained from the literature search was used to investigate some preliminary experiment design considerations for a small angle neutron scattering experiment on LCPs. In addition, a survey of world wide neutron scattering facilities was made. Both their physical capabilities and their user requirements were determined. The results of the survey and the preliminary experiment design considerations were used to recommend a primary and a first alternate neutron scattering facility at which to perform the planned experiments.

PREPARATIONS FOR NEUTRON SCATTERING INVESTIGATIONS
OF LIQUID-CRYSTAL POLYMERS

David M. Elliott

INTRODUCTION

The development of advanced new materials possessing very high specific strength and specific modulus is critically important to the Air Force mission in both the exploration of interplanetary space and the development of higher performance and lower cost missile systems. Composite materials with superior mechanical properties have been made from high-strength fibers and graphite. Parts and structures made from these composite materials are, however, relatively costly because of the complexity of the manufacturing process. The high-strength fibers used in these composite materials are prepared from stiff-chain aromatic polymers. These polymers are called stiff-chain polymers because the molecules tend to be quite elongated in solution or in the melt. These elongated molecules have a propensity to form a definite crystalline organization in the liquid state (either solution or melt) and consequently they are referred to as liquid-crystal polymers (LCPs). In addition to high-strength fibers, LCPs can be injection or blow molded inexpensively into complex shapes. The liquid crystalline order persists and is often enhanced as LCPs are carried into the solid state. Solid materials made from LCPs are, however, inhomogeneous and highly anisotropic. They can be thought of as molecular composites.

The mechanical properties of solid LCPs depend importantly on the processing as well as on the material itself. Flow, either in the spinning of fibers or in the molding of parts, can result in an increase in the crystalline order and this in turn can result in significant enhancements of both strength and modulus. Temperature is also an important processing variable. Some LCPs are known to undergo very significant and favorable changes in both mechanical and chemical properties after thermal annealing. The relationships between molecular conformation and crystalline order and mechanical and chemical properties are not well known yet understanding these relationships is vital to developing these molecular composites to their full potential. Recognizing this, researchers at the Phillips Laboratory are pursuing a

program of both theoretical and experimental studies to characterize the molecular structure, dynamics, and mechanism of formation of the annealed state in selected linear LCPs. The experimental part of the program necessarily involves a number of different techniques including thermal and mechanical as well as atomic (X-ray) and nuclear (neutron scattering).

The nuclear techniques will include as a minimum small angle neutron scattering and neutron diffraction and may include neutron reflectometry as well. The neutron scattering and diffraction experiments being considered for this program will require very high integrated neutron fluxes and complex experimental facilities which are available only at large centrally located facilities. There are very few facilities in the world where these experiments can be carried out, probably less than twenty five in total. The neutrons are produced either by high-energy positive ion accelerators or by nuclear research reactors. The accelerator sources produce higher instantaneous neutron fluxes (and higher neutron energies) but their beams are pulsed resulting in somewhat lower integrated fluxes. The facilities based on accelerator sources have historically concentrated on pulsed source instrumentation and techniques. The development of equipment and techniques for the exploration of basic issues in condensed-matter science has taken place largely at research reactor based facilities. For these reasons, only research reactor based neutron scattering facilities are being considered for the Phillips Laboratory LCP experimental research.

The objectives of the present work are all related to preparations for the neutron scattering investigations to be carried out in the immediate future. In close cooperation with other researchers at the Phillips Laboratory, the following principal tasks were initiated:

1. Preliminary search of the literature for information about SANS and liquid crystal polymers.
2. Preliminary design of SANS experiments.
3. Survey of world-wide neutron scattering facilities appropriate for the Phillips Laboratory LCP experimental work and selection of a primary and first alternate facility.

NEUTRON SCATTERING AND APPLICATIONS TO LCPs

Neutrons are electrically neutral sub-atomic particles with a mass of $1.675 \text{ E-}27\text{kg}$, spin $1/2$, and a magnetic moment of 1.913 nuclear magnetrons. The de Broglie relation, $\mathbf{p} = \hbar\mathbf{k}/2\pi$, is the basis for conversions among neutron properties expressed in terms of momentum (\mathbf{p}), velocity ($V_n = \mathbf{p}/m_n$), wave vector (\mathbf{k}), wavelength (λ), and kinetic energy (E). Table 1. (Reference 1) gives some useful conversion relations.

Energy E (meV)		1	10	50
corresponds to	defined by			
Angular frequency ω (rad/sec)	$\omega = E/\hbar$	1.52×10^{12}	1.52×10^{13}	7.6×10^{14}
Frequency ν (Hz)	$\nu = E/h$	2.42×10^{11}	2.42×10^{12}	1.21×10^{13}
Optical wave number σ (cm^{-1})	$\sigma = \nu/c$	3.065	30.65	403.3
Temperature (T , K)	$T = E/k_B$	11.605	116	580
Neutron velocity v_n (m/sec)	$v_n = (2E/m_n)^{1/2}$	437	1383	3093
Neutron wavelength λ (Å)	$\lambda = h/(2m_n E)^{1/2}$	9.04	2.86	1.28
Neutron wave number k (Å^{-1})	$k = (2m_n E)^{1/2}/h$	0.695	2.2	4.91

Table 1. Neutron Conversion Data

Neutrons in thermal equilibrium with their surroundings (called "thermal neutrons") turn out to be an extremely useful probe of the structure of condensed matter. This fact was recognized by physicists soon after the discovery of the neutron in 1932. Table 2. (Reference 2) shows some characteristics of neutrons at selected energies.

Quantity	Unit	Definition	Ultracold	Cold	Thermal	Epithermal
Energy E	meV ^a		0.00025	1	25	1000
Temperature T	K	E/k_B	0.0029	12	290	12,000
Wavelength λ^b	Å	$h/(2mE)^{1/2}$	570	9.0	1.8	0.29
Wave vector k^c	Å ⁻¹	$(2mE)^{1/2}/h$	0.011	0.7	3.5	22
Velocity v^d	m/s	$(2E/m)^{1/2}$	6.9	440	2200	14,000

Table 2. Characteristics of Neutrons at Selected Energies

Some of the reasons why neutrons are such useful probes of condensed matter listed below. (Reference 1)

1. The interaction of thermal neutrons with matter is weak.
2. The absorption of thermal neutrons by matter is often very small.
3. The wavelength of thermal neutrons is comparable to atomic and molecular spacing.
4. The nonmagnetic scattering amplitude of neutrons does not depend on the electronic structure, rather it depends on the nuclear structure. It is often very different for different isotopes of the same element.

The weak interaction and small absorption of neutrons allows them to be used to probe at much greater depths into a material than electromagnetic radiation (X-rays or light). The ability to probe at depth is critical to the current LCP research because the molecular orientation and perhaps the conformation as well are strong functions of position within the material and particularly near the surface.

Neutron scattering is a relatively new technique of material science, having been extensively utilized only over the last fifteen years or so. A considerable amount of fundamental research using neutron scattering techniques has been done on polymers and the experimental techniques are relatively well developed. The picture is, however, not as bright for LCPs. Very little neutron scattering work has been done with LCPs. The majority of the work that has been done in this area is very recent and was done in Europe. (References 7-10)

Neutron scattering can, in general, be reduced to the simple form: (Reference 15)

$$\frac{d^2\sigma}{d\Omega dE} = AS_{\alpha\beta}(Q, \omega)$$

Where: The measured intensity is proportional to
(perhaps integrated over E)

$$\frac{d^2\sigma}{d\Omega dE}$$

A is a (known) constant which contains all of the neutron properties

$S_{\alpha\beta}(\mathbf{Q}, \omega)$ is determined solely by the sample.

$$\mathbf{Q} = \mathbf{k}_1 - \mathbf{k}_2$$

$$E = \hbar \omega = (\hbar/2m) * (\mathbf{k}_1^2 - \mathbf{k}_2^2)$$

and α, β label the spin state of the neutron before and after the scattering event.

$k = 2\pi/\lambda = mv/\hbar$ and \mathbf{k}_1 and \mathbf{k}_2 are the neutron wave vectors before and after the scattering event respectively.

Thus a neutron scattering experiment reduces to a determination of the magnitude and direction \mathbf{k} both before and after the scattering event.

In the experiments of interest for the current program, the scattering is elastic. The magnitude of the neutron wave vector is thus unchanged during the scattering event. For these experiments, one needs only to measure the magnitude of the wave vector (or the energy) once, usually before the scattering event. One then measures the direction of the emergent neutrons after the scattering event. These experiments are simple in theory. The difficulties, as usual, lie in the details. It is the details that must be determined in the design of specific experiments.

RESULTS

Preliminary Literature Search

The experimental phase of the Phillips Laboratory LCP Research Program contemplates using a number of different neutron scattering techniques. As mentioned previously, these techniques include neutron diffraction (ND), neutron reflectometry (NR), and small angle neutron scattering (SANS). The SANS technique produces a wider variety of useful information than do the others. Not surprisingly, SANS is the most demanding of these techniques. SANS requires a much greater integrated neutron flux, better collimation and wavelength selectivity, a more massive, complicated, and higher precision spectrometer, and a position sensitive detector. Perhaps because they are so complicated and expensive, there is more variability in the capabilities of

the SANS instruments at different scattering research facilities than in the instruments used for the other techniques. For these reasons and because the principal objective of the author's summer work was to recommend a primary and a first alternate reactor based neutron scattering facility, the literature search began with and still concentrates on small angle neutron scattering.

The essential objective of the SANS literature search is to locate and obtain copies of references that will provide or support the development of:

1. A thorough understanding of the theory of the various possible scattering interactions, particularly SANS.
2. The ability to model the contemplated SANS experiments and predict specific results based on knowledge of the experimental conditions and of the scattering medium.
3. A knowledge of the techniques used and the specific results obtained for polymer systems in general and LCPs in particular.

During the summer, copies of a number of good references were obtained. Three chapters in two books and one journal article were obtained which provide an excellent introduction to the theory of neutron scattering. (References 1-4) Five references dealing with the application of SANS to polymer systems (References 5-9) and five dealing with specific applications to liquid crystal polymers (References 10-13, 15) were obtained. These references have already proven their worth in providing information useful in the preliminary design of SANS experiments.

Preliminary SANS Experiment Design Considerations

One important purpose of the initial preliminary SANS experiment design was to secure sufficient information to assess the suitability of a number of candidate neutron scattering facilities for performing the planned LCP experiments. The SANS experiment design was used for this purpose because it is more demanding than ND and because virtually every neutron scattering facility has adequate ND capabilities. The SANS technique is extremely versatile and a number of different parameters of interest can be measured.

Two Q regions of interest to the polymer scientist are often identified within the small angle range. (Reference 8) The first and lowest region is known as the Guinier region and is defined by $1/Q > R_c$ where R_c is the radius

of gyration or the second moment of the distribution of the scattering centers. In this region, the scattering is sensitive to the dimensions of the chains and not their detailed shape. The second region is defined by $1/R_g < Q < 1$ where l is the length of a characteristic segment of the chain, usually a few monomer units. In the second region, the scattering becomes very sensitive to correlations between segments and thus to the shape of the molecule. In this region, $S(Q)$ varies as $1/Q$ for a rod like structure and as $1/Q^3$ for a sphere.

The first experiment considered was solid state molecular weight determination for amorphous LCPs. This experiment utilizes a variation of the Zimm plot technique. (Reference 7) Here Q is approximately in the Guinier region and the desired information is obtained from a plot of the inverse of the scattered intensity as a function of the square of Q . Since the Guinier region is defined as $1/Q > R_g$, the high end of the Q range is roughly $1/R_g$ and the low end is as low as the instrument will allow. In the literature, these Q ranges usually extend over a factor of from four to ten. (References 10,11,12) The materials of interest for the current LCP research are rod like molecules with a length on the order of 90 \AA . This would suggest a radius of gyration of about 26 \AA . The corresponding Guinier range (in Q) for these molecules would be about 0.005 to 0.04 \AA^{-1} . This then was taken to be the Q range that had to be available. The low end of this range is at about the limit of some SANS machines. Most machines will cover up to at least 0.4 and some will go up as high as 1.0 .

Given that the Q range is acceptable, probably the next most important factor is the neutron flux available on the sample at the required wavelengths. A number of things affect the available flux. The most obvious of these are the peak thermal flux available in the reactor and whether or not a cold moderator (cold source) is present. The peak thermal flux available in the high power reactors ($60 - 100 \text{ MW}$) is about $10^{15} \text{ n/cm}^2/\text{sec}$. The medium power reactors ($10 - 20 \text{ MW}$) have peak thermal fluxes on the order of 10^{14} . The addition of a cold moderator will increase the flux of neutrons of interest in SANS experiments by about a factor of 5. The flux available at the sample is a very important factor in determining the time required to perform a given experiment (counting time for adequate statistics). For the best combination of a high power reactor with a cold moderator, counting times for typical SANS experiments with polymers vary from about 5 to about 60 minutes for isotropic scattering and from about 25 to about 300 minutes for anisotropic scattering (as in ordered polymer systems). (Reference 7) Without using any tricks to optimize the experimental setup, these times would increase by about a factor of 50 for a medium power reactor without a cold

moderator. Fortunately, there are a number of tricks that can be used to increase the total number of neutrons being scattered from the sample as well as those impinging upon it.

The large difference in scattering length between deuterium and hydrogen provides excellent contrast between molecules labeled with deuterium and those that are not. Increasing the concentration of deuterated chains up to a maximum of 50 percent or so increases the signal to background ratio. On the other hand, deuteration is not without its potential drawbacks. One major drawback is the relatively high cost. Another is the possibility that deuteration may affect the results of the experiment in subtle and unpredictable ways. "One of the basic assumptions of SANS is that deuterium labeling does not affect the properties of the chains. In the case of PE, deuterium labeling causes a 6° difference in the melting temperature, and thus during slow-melt crystallization a nonrandom distribution of deuterated chains occurs. (Reference 7) "The SANS method is extremely sensitive to small deviations from a statistical distribution. (Reference 5)

The size of the sample, of course, impacts the number of neutrons reaching the detector. In general, the sample diameter should be as large as possible within the confines of the beam geometry at the sample. The sample thickness will be governed by the specific material under investigation and the concentration of deuterated chains. The literature provides some useful guidance here. "Small angle scattering is usually carried out with samples that have 50% transmission. A mainly hydrogenous sample will be about 1 mm thick, whereas a sample that contains a high proportion of deuterium can be 5-10 mm thick." (Reference 7)

Neutron Scattering Facility Survey

One important question that needed to be answered early in the program was: Which neutron scattering facilities would be likely to have both the required physical resources and a user policy that would allow the planned experiments to be carried out in a timely and cost effective manner. The Phillips Laboratory program plan calls for the neutron scattering experiments to be done over a relatively narrow time frame beginning in the very near future. To complicate matters even further, the program must operate under very tight budgetary constraints. To begin the process of answering this question, a list of all of the major reactor based neutron scattering facilities in the world was made. Entries on this list were obtained from the literature and a from a series of personal contacts and telephone calls.

Table 3. lists all of the major facilities currently known to the author.

All of the facilities listed in Table 3. have a number of different neutron scattering instruments and all have at least one SANS spectrometer. If any facility could not meet the physical requirements of the proposed experiments, it would most likely be because of limitations of the SANS instrument. Preliminary experiment design considerations suggest that the most probable limiting factors are the Q range and the neutron flux available on the sample. It turns out that any of the facilities listed in Table 3. would be physically capable of performing all of the planned experiments. The main physical difference among these facilities would be the time it takes to perform each individual experiment. The high power reactors with a cold moderator can do the experiments in much less time.

Having developed rapidly over the past fifteen years or so, neutron scattering is now seen as a core competency in materials science and engineering. From Table 3. one can see that the total number of neutron scattering facilities in the world is very small. The result is that most of the world's neutron scattering facilities are over subscribed, some by as much as a factor of three. There are simply more good experiments to be run than there are facilities to run them. As a result, most of these facilities have adopted policies that restrict the number of experiments that can be run. Many of them have adopted the "user facility" approach where instrument time is allocated by a committee based on their evaluation of the technical merit of proposals submitted by those who want to use the facility. There are a number of drawbacks to this approach from the experimenter's point of view. The proposal review and time allocation process takes a long time, often several months. Since all proposals cannot be approved, there is a substantial probability that any individual proposal might not be

United States

- o Brookhaven National Laboratory (HFBR)
Upton, Long Island New York
- o Oak Ridge National Laboratory (HFIR)
Oak Ridge, Tennessee
- o National Institute of Standards (NIST)
Gaithersburg, Maryland
- o University of Missouri (MURR)
Columbia, Missouri

Australia

- o Australian Nuclear Science & Technology Organization (ANSTO)
Lucas Heights

France

- o Institut Laue Langevin (ILL)
Grenoble
- o Orphee Reactor Center (ORPHEE)
Saclay
- o SILOE, CENG (SILOE)
Grenoble

Germany

- o KFA Research Reactor (JULICH)
Julich
- o HMI Reactor Facility (HMI)
Berlin

Japan

- o JAERI III Reactor Facility (JAERI)

Table 3. Worldwide Reactor Based Neutron Scattering Facilities

accepted at all. Many proposals end up in an alternate position or in a queue. The time and budgetary constraints of the Phillips Laboratory LCP experimental program are such that it would be very difficult to do all of the proposed experiments at a user facility, simply because of the process required to obtain instrument time.

From those facilities listed in Table 3, five candidate facilities at which the experimental program might reasonably be carried out in a timely manner were selected. Four of these are in the United States and one is in Australia. Next, detailed technical and user policy information was collected for each candidate facility. Literature references were reviewed and printed materials were obtained where these were available. Personal contacts were made by telephone at each candidate facility. A summary of the information is given in Table 4.

	<u>HFBR</u>	<u>HFIR</u>	<u>HIFAR</u>	<u>NIST</u>	<u>MURR</u>
Reactor Power (MW)	60	100	10	20	10
Peak Thermal Flux (E14)	5	11	1.6	4	1
Cold Moderator?	Yes	No	No	Yes	No
Number of Spectrometers	13	9	8	12	8
User Facility?	Yes	Yes	No	Yes	No

Table 4. Features and User Policy of Candidate Facilities

All of these candidate facilities are (or soon will be) physically capable of providing the adequate Q range, neutron flux and other capabilities required to carry out the Phillips Laboratory LCP experimental program. Only two of these facilities, however, are not user facilities and thus would be capable under present circumstances of meeting all of the current program requirements: HIFAR and MURR. Both of these facilities have user policies that allow and, in fact, encourage collaboration between their own scientists and those of the experimenter. Formal proposals are not required at either of them. The reactor power and peak flux are similar and the SANS spectrometers are nearly identical. Both SANS spectrometers are currently being built or upgraded. Present schedules call for the SANS spectrometer at HIFAR to be completed in mid to late 1992 while that at MURR is planned for operation in the first or second quarter of 1993. Based on all of the above factors, it is recommended that the HIFAR facility be selected as the primary facility and MURR as the first alternate.

CONCLUSIONS

Neutron scattering techniques, particularly small angle neutron scattering and neutron diffraction, are capable of providing significant insight into both the structure and the morphology of liquid crystal polymers in the solid state. These techniques should be uniquely useful in helping to understand the molecular structure, dynamics and mechanism of formation of the annealed state. There are, however, significant problems in actually applying these tools to LCPs. These problems lie in two primary areas: The first area is the lack of previous work using neutron scattering techniques specifically for investigating LCPs. Only a few publications are available and these address only a small part of the total picture. The theory is adequately presented in the literature but there is little guidance for the actual application of the techniques and the interpretation of the results. The second area is the difficulty of actually getting instrument time at one of the very few neutron scattering facilities in the world capable of doing the work.

The essential conclusions reached from this work, then, can be summarized as follows:

1. Neutron scattering techniques, particularly SANS and ND are inherently capable of answering many of the fundamental structural questions of interest to the Phillips Laboratory LCP Research Program.
2. A very considerable effort is required before the actual experimental program can begin. The primary things that need to be done are a complete literature search and detailed experiment design.
3. Considering all of the requirements of the Phillips Laboratory program, particularly temporal and budgetary constraints, the preferred neutron scattering facility at which to do the work is the HIFAR facility in Australia. A close second is the MURR facility in Missouri.

RECOMMENDATIONS

The following two recommendations are made based on the results the Summer Research Program and discussions with the Phillips Laboratory research scientists and others in the neutron scattering community:

1. Complete the SANS literature search and extend it to include ND. The essential objective of the literature search should be to locate and obtain copies of references that will provide or support the development of:
 - A. A thorough understanding of the theory of the various possible scattering interactions.
 - B. The ability to model the contemplated experiments and predict specific results based on knowledge of the experimental conditions and of the scattering medium.
 - C. A knowledge of the techniques used and the specific results obtained for polymer systems in general and LCPs in particular.
2. Complete the preliminary experiment designs done in the current work and extend them to include ND.

These recommendations are included in the objectives of the author's AFOSR Research Initiation Proposal.

REFERENCES

1. *Neutron Scattering--General Introduction* by G. Kostorz and S.W. Lovesey in *Treatise on Materials Science and Technology*, Volume 15, Neutron Scattering, Academic Press, New York, 1979
2. *Introduction to Neutron Scattering* by D.L. Price and K. Sköld in *Methods of Experimental Physics*, Volume 23, Neutron Scattering, Part A, Academic Press, New York, 1986
3. *Small-Angle Scattering and its Applications to Materials Science* by G. Kostorz in *Treatise on Materials Science and Technology*, Volume 15, Neutron Scattering, Academic Press, New York, 1979
4. *Neutron Scattering and Amorphous Polymers*, A. Maconnachie and R.W. Richards, *Polymer*, Vol 19, July, 1978
5. *Small-Angle Neutron Scattering Studies of Molecular Clustering in Mixtures of Polyethylene and Deuterated Polyethylene*, J.Schelten et al, *Polymer*, Vol 18, November, 1977
6. *Chain Conformation in Molten Polyethylene by Low Angle Neutron Scattering*, J. Schelten et al, *Letters in Polymer*, Vol 15, October, 1974
7. *Polymers* by J.S. Higgins and A. Maconnachie in *Methods of Experimental Physics*, Volume 23, Neutron Scattering, Part C, Academic Press, New York, NY, 1987
8. *Polymer Conformation and Dynamics* by J.S. Higgins in *Treatise on Materials Science and Technology*, Volume 15, Neutron Scattering, Academic Press, New York, NY, 1979
9. *Small-Angle Neutron Scattering Measurement of Block Copolymer Interphase Structure*, R.W. Richards and J.L. Thomason, *Polymer*, Vol 24, September, 1983
10. *Conformational Anisotropy of Liquid Crystalline Side Chain Polymers: A Small Angle Neutron Scattering Study*, F. Moussa et al, *Le Journal De Physique*, 48, 1987, pp 1079-1083

11. *Liquid Crystal Polymers: A Small Angle Neutron Scattering Study*, F. Hardouin et al, *Molecular Crystals and Liquid Crystals*, Vol 155, 1988, pp 389-397
12. *Side Chain Mesomorphic Polymers: Studies of Labelled Backbones by Neutron Scattering*, P. Keller et al, *J. Physique Letters*, 46, 1985, pp L-1065-L-1071
13. *Small Angle Neutron Scattering Study of Comb-Like Liquid Crystal Polysiloxane Macromolecule Conformation in Smectic Phase*, J. Kalus et al, *Molecular Crystals and Liquid Crystals*, Vol 155, 1988, pp 347-356
14. *Major Facilities for Materials Research and Related Disciplines*, National Research Council, PB85-108850, Washington, DC, July, 1984
15. *Liquid Crystals* by J.A. Janik and T. Riste in *Methods of Experimental Physics, Volume 23, Neutron Scattering, Part B*, Academic Press, New York, NY, 1987
16. *Neutron Properties and Production*, J.M. Rowe, *Workshop on Neutron and Synchrotron Radiation Studies of Materials*, Brookhaven National Laboratory, Upton, NY, May, 1992
17. *Neutron Probes Tackle Industrial Problems*, H.J. Prask, *Advanced Materials & Processes*, 9/91
18. *Industrial Applications* by M.T. Hutchings and C.G. Windsor in *Methods of Experimental Physics, Volume 23, Neutron Scattering, Part C*, Academic Press, New York, NY, 1987

ESTABLISHMENT OF AN ARCJET OPTICAL
DIAGNOSTICS FACILITY AT PHILLIPS LAB

Daniel A. Erwin
Associate Professor
Department of Aerospace Engineering

and

John H. Schilling and Jeff A. Pobst
Research Assistants
Department of Aerospace Engineering

University of Southern California
University Park
Los Angeles, CA 90089

Final Report For:
Summer Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

August, 1992

ESTABLISHMENT OF AN ARCJET OPTICAL
DIAGNOSTICS FACILITY AT PHILLIPS LAB

Daniel A. Erwin, John H. Schilling, and Jeff A. Pobst
Department of Aerospace Engineering
University of Southern California

Abstract

An arcjet optical diagnostic facility was established at the Phillips Laboratory. This facility is expected to provide accurate measures of flow properties in the plume and nozzle region of arcjet thrusters to assist in the development of more efficient arcjet thrusters for spacecraft stationkeeping and orbit-raising applications. A 30-kW arcjet was mounted and successfully operated in a vacuum test chamber capable of accurately duplicating the space environment. Optical elements were installed to allow active or passive measurement of flow properties using the techniques of emission spectroscopy and laser-induced fluorescence. A tunable ring dye laser pumped by a 20-watt CW argon-ion laser were installed and aligned for LIF experiments. Finally, an automated control and data-acquisition system was installed and software developed to conduct a variety of experiments.

ESTABLISHMENT OF AN ARCJET OPTICAL
DIAGNOSTICS FACILITY AT PHILLIPS LAB

Dan Erwin, John Schilling, and Jeff Pobst

INTRODUCTION

Arcjets are a type of electrothermal rocket in which a gaseous propellant is heated by a high-current electric arc and expanded through a nozzle to produce thrust. Such thrusters have a specific impulse in the 800 to 1500 second range, compared with 300-450 seconds for chemical rockets. Although such thrusters do not produce enough thrust for launch vehicle applications, they promise significant advantages when used for spacecraft stationkeeping and orbit-raising purposes. In particular, the reduced propellant mass due to increased specific impulse will allow an increase in payload or a reduction in launch vehicle size and cost. However, current arcjet thrusters are remarkably inefficient (~30%) in electric power utilization, so further research is necessary before arcjets can be used on operational spacecraft. In order to conduct such research, it is necessary to measure flow properties such as velocity and density in the arcjet nozzle and plume. This will increase our understanding of the processes affecting arcjet efficiency as well as serving to validate theoretical and computational models of arcjet performance. In addition, if the experimental process is simple enough, it can be used to rapidly evaluate a large number of proposed arcjet designs.

THE ARCJET

The test facility is designed to be compatible with virtually any arcjet of up to 30kW power. Initial tests were conducted using a 30kW arcjet designed and manufactured by the Air Force Astronautics Lab. Both the cathode and anode/nozzle are machined from thoriated tungsten due to the extreme temperatures produced by the arc, and boron nitride is used as an insulator for the same reason. Although the arcjet is capable of operating on a wide variety of propellants, all tests in this facility will be using gaseous ammonia, for which the appropriate flow rate is 0.25gm/sec. The arcjet requires approximately 300 amps at a voltage of 100 volts, for a total of 30kW. A brief pulse of 1-2kV is normally required for arc initiation, but by using argon propellant at a lower flow rate the arc can be initiated at ~300V.

The arcjet is expected to operate at a specific impulse of approximately 800s on ammonia propellant, with a thrust at maximum power of approximately two newtons. This corresponds to a thrust power of 8kW, giving an efficiency of 25-30%. As noted above, the purpose of this facility is to assist in

efforts to increase this figure, hopefully to 50-65%. It is believed that most of the wasted energy goes to thermal, boundary layer, and frozen-flow losses, but the precise mechanisms are not well understood.

ARCJET TEST CHAMBER

Arcjet engines are intended solely for space applications, and have a very low nozzle exit pressure. Thus, it is necessary to test them in a vacuum environment. Previous experience suggests that a pressure of less than 100 milli Torr is desirable. An aluminum vacuum chamber approximately 1.75 meters in diameter and 3.5 meters long was provided for this purpose, connected by large-diameter steel ducts to a Roots-Stoke vacuum pump combination capable of a maximum 10,800 cfm pumping rate. Given a 0.25gm/second propellant flow rate into the vacuum chamber, this plumbing system should be able to maintain a pressure of 60 milli Torr in the test chamber. However, due to design and manufacturing flaws, the vacuum chamber suffered major leaks and could not be pumped below 400 milli Torr, even with no propellant flow. By replacing numerous O-ring seals and redesigning several flanges, the desired performance was achieved.

The heat load caused by the operation of a 30kW arcjet can destroy the seals of the vacuum system, and could pose a safety hazard to people working around the vacuum chamber. Thus, a cooling system is necessary. A cooling jacket was provided with the vacuum chamber, consisting of a thin cylindrical liner with a network of water channels of approximately 2 cm diameter. It is mounted on supports maintaining a uniform 10 cm separation between tank wall and liner. A cooling panel of similar design was mounted on the end of the test chamber. Some plumbing work was necessary to connect the cooling panels to a water pumping and refrigeration system.

The arcjet was mounted on a test stand in the center of the vacuum chamber. Two test stands are available for this purpose. The first is a simple vertical post with a mounting bracket for the arcjet and fittings to attach the test stand to the interior of the chamber. This served to hold the arcjet for the initial facility tests. The arcjet mounting bracket is a simple clamp, so no specific mating flange on the arcjet is required. This is consistent with the desire that the test facility be compatible with a wide variety of arcjets. A more sophisticated thrust stand was mounted on a multi-axis translation stage, allowing the arcjet to be positioned in three dimensions with a high degree of accuracy. This was necessary to probe different points in the flow field, as the optical path of the diagnostic equipment was fixed. Motive power was provided by three high-precision

stepper motors under computer control. This allowed automated positioning of the arcjet without opening the chamber.

PROPELLANT FEED

No propellant feed system was provided for the test chamber. 150-lb bottles of liquid anhydrous ammonia were stored at a location outside of the laboratory due to safety concerns. A 0.25" stainless-steel propellant line was run between the storage bottle and the vacuum chamber, with a bulkhead feed-through allowing the propellant to enter the test chamber. Due to the motion of the arcjet on the test stand, a flexible braided-steel propellant line was necessary to feed propellant to the arcjet from this point. Immediately before entering the arcjet, the propellant passed through a length of machined alumina tubing to provide electrical isolation of the arcjet.

The ammonia storage bottle used to provide propellant was mounted on a high-precision electronic scale to monitor ammonia consumption. A regulator manifold was used to provide a constant 80-psi supply pressure as well as several auxiliary capabilities, including emergency pressure relief and nitrogen purge of the propellant feed system. A tee in the propellant line at this point allowed an argon bottle to be attached in parallel with the ammonia, as the use of argon propellant allows a lower voltage to be used in the arc-initiation process. Once the arcjet is running smoothly, the propellant is switched from argon to ammonia.

A Sierra flow controller was used to regulate the propellant mass flow rate. This type of controller uses a thermocouple to monitor convective heat transfer rates, which are directly proportional to mass flow rate for a given fluid. A needle valve with a feedback control system maintains the flow rate at the desired level. A four-liter plenum was placed directly downstream of the flow controller to prevent rapid fluctuation of the flow rate, as catastrophic damage can occur to the arcjet is operated with too low a flow rate.

POWER SUPPLY

Operational arcjets are expected to be powered by high-frequency switching power supplies, due to their light weight and high efficiency (~95%). However, such power supplies are complex and expensive, and where weight and power limitations do not apply, simpler systems are acceptable. As an arcjet is electrically quite similar to an arc welder, an L-Tec high-current arc welding power supply was selected. Although this is

nominally a current-regulated power supply, the negative voltage/current characteristic of an arcjet would lead to unacceptable instabilities in supply voltage and current. Thus, a ballast resistor must be placed in series with the arcjet.

For this application a 110m Ω resistor was fabricated and mounted in an insulated enclosure for safety. The power supply and ballast resistor were connected by insulated cable to a set of insulated bulkhead feed-throughs on the test chamber. Inside the chamber, a flexible welding cable was used to provide power to the arcjet without interfering with the freedom of motion of the arcjet. A separate high-voltage starting circuit was considered, but it was decided that by using argon propellant for starting, no separate circuit was necessary. Arc initiation and current control are provided by the front-panel controls of the L-Tec power supply.

One difference between the arc-welding power supply and a high-frequency switching power supply is that the latter has a current ripple of ~10-15% at the switching frequency, usually about 10 KHz. This has been shown to affect the flow field and cathode erosion rates of an arcjet. It is not possible to study these effects using a ripple-free welding power supply, but it is expected that these effects will not significantly change the performance of the arcjet and that the test results will still be valid. Nonetheless, access to a switching power supply at a later date would be desirable.

LASER DIAGNOSTIC SYSTEM

Laser Induced Fluorescence (LIF) is a non-intrusive technique for determining velocity and temperature profiles in arcjet plumes. To excite plume gasses with laser radiation one must be able to tune the frequency of the laser light to that corresponding to the desired energy level transition of the probed gas (here, the Balmer alpha transition of atomic hydrogen). When the gas is moving toward the incoming laser light, the frequency where excitation will occur is doppler shifted in relation to the velocity of the gas. By knowing the frequency where the gas should absorb the laser light, the difference in frequency where the gas does absorb the laser energy and fluoresces relates directly to the velocity the gas is traveling and a velocity measurement may be made. The equation $v = \frac{c(\lambda - \lambda_0)}{\lambda_0}$ describes the relation where c is the speed of light, λ_0 is the expected transition wavelength, λ is the measured transition wavelength, and v is the velocity.

In order to use this technique, the laser must be freely tunable over a range of frequencies. Furthermore, the linewidth must be quite narrow. As quantum effects prevent pulsed lasers from achieving narrow line widths, a continuous wave (CW) laser must be used. The tunability requirement is satisfied by a dye laser, which does not depend on any particular atomic transition and can therefore operate at any wavelength within a specified range. , A Coherent 899 CW tunable ring dye laser system was particularly chosen for its wide tunability range ($\sim 10 \text{ \AA}$) and narrow linewidth ($\sim 0.0003 \text{ \AA}$) which allows a precise velocity measurement over a tuning range large enough to record the full lineshape so that broadening mechanisms can be fully measured. A dye laser requires a pump source, usually a laser operating at a higher frequency than the dye laser. A 20W argon-ion laser was borrowed from the High Energy Density Materials research group at Phillips Lab for this purpose.

The Coherent 899 is shipped with a dedicated computer for frequency control. This computer was slaved to the facility control/data acquisition computer to allow automated tuning and scanning of the laser frequency. Without this ability, even the simplest of experiments would have taken many hours as the laser was manually retuned literally hundreds of times. With the autoscan feature, the full line shape could be measured in a matter of minutes.

In order to fully characterize the plume flow, it is necessary to measure both longitudinal and radial velocities (it is assumed that the flow is radially symmetric). As the LIF technique measures velocity in the direction parallel to the beam, two different beam paths were necessary. Thus, the beam is split by a prism beamsplitter immediately after leaving the laser. One beam is directed horizontally into the end port of the test chamber and used to measure longitudinal velocity, the other is directed under the chamber where it enters vertically through a small window on the bottom of the test chamber. In both cases, prisms were used to direct the beam, as the high power densities involved (millions of watts per square meter) would have destroyed the coatings of optical mirrors. A system of lenses was used to insure that the beam diameter was less than one millimeter at the test area in order to probe the smallest possible volume of the plume.

The amount of laser energy absorbed by the plume cannot easily be measured directly. However, most of the absorbed energy is reradiated as a fluorescence signal, which can be detected. For this purpose, a photomultiplier tube (PMT) was selected. Although a PMT has no imaging capability, it can accurately measure the intensity of very weak light signals with high accuracy and time resolution. The PMT was attached to a housing on the side port of the test chamber. In this position, the collection path is perpendicular to both laser paths, simplifying interpretation of the

experimental results. A set of lenses was used to insure that only fluorescence from a specific point in the plume was collected. This, combined with the narrow laser beam diameter mentioned earlier, insures a high degree of spatial resolution of the measurements. Notch filters were used to limit the signal to that produced at or near the desired frequency.

Because the fluorescence signal is small compared to the thermal emission of the plume, it was necessary to filter out the background noise. A lock-in amplifier was selected for this purpose. By periodically interrupting the laser beam, it compares the total signal with the background noise. Over a period of time, the desired signal can be isolated and amplified with a high degree of accuracy despite the small signal-to-noise ratio.

CONTROL AND DATA ACQUISITION

The experiment was controlled and the data acquired by a Unisys computer using an Intel 386/20MHz processor. Although not specifically designed for such applications, this is a very versatile system, as its compatibility with the IBM PC line allows the use of a wide variety of commercial software and accessories. General Purpose Interface Bus (GPIB) and Analog-to-Digital conversion boards from National Instruments were installed to enable the computer to control the other equipment used in the experiment as well as to monitor several sensors, including the PMT.

This hardware had the capability to run the test facility with minimal human intervention. In order to achieve this goal, appropriate software had to be used. The Borland C++ software development kit, along with application libraries provided by National Instruments, was used to write and compile a program capable of fully automatic control of the experiment. Much of the software development was done under contract at the University of Southern California, where a similar computer control/data acquisition system is used in arcjet diagnostic experiments.

LASER INDUCED FLUORESCENCE EXPERIMENTS

The primary purpose of this facility is to use the laser-induced fluorescence technique to characterize the plume flow of any arcjet. As described earlier, this involves measuring the fluorescence produced as laser light is absorbed by the plume gasses. By measuring the frequency of maximum absorption and comparing to the known frequency of the atomic transition being excited, the doppler shift and thus the velocity of the plume can be determined. Also, by measuring the shape of the absorption profile, the magnitude of doppler broadening and stark broadening can be determined. These are proportional to kinetic temperature and electron density, respectively.

Finally, the intensity of the fluorescence signal may be used to measure the density of the particular species and excited state being probed.

In order to obtain this data, the arcjet is positioned by the traverse system so that the intersection of the laser beam and the PMT collection path occurs at the desired point in the plume. The laser is then activate and tuned to a wavelength several angstroms below that of the desired transition, and the PMT signal, amplified by the lock-in amplifier to eliminate background noise, is recorded. The laser wavelength is then increased by a small amount, and the process is repeated. This occurs until the wavelength is several angstroms above the desired transition, and an entire absorption profile has been measured and recorded. At this point, the engine is repositioned and another absorption profile is measured. Ultimately, the entire plume will be characterized, with longitudinal velocity, radial velocity, temperature, electron density, and excited state density measured with high accuracy and sub-millimeter spatial resolution.

As the entire process, including data reduction, is automated, the experiment takes only a matter of hours. Thus, new arcjet designs or modified nozzle configurations can be easily tested. This will allow rapid verification of new designs or of theoretical and computational models of arcjet operation. Such a capability should be quite useful in accelerating the pace of arcjet research.

Due to the fact that the laser and collection paths must be perpendicular to achieve high spatial resolution and accurate doppler measurements, it is not possible to measure flow properties in the nozzle using a standard arcjet. While complete knowledge of the plume flow may give a great deal of insight into the nozzle flow, it may be necessary to directly measure nozzle flow in this case. It is possible to modify an arcjet to allow this. If a narrow vertical slit is machined in the arcjet nozzle, and a properly contoured quartz window can be inserted in the slit, the arcjet can be operated without significant disruption of the nozzle flow. If the laser is directed longitudinally down the nozzle and the collection path is through the window, the nozzle flow can be measured and characterized in the plane of the window.

EMISSION SPECTROSCOPY

Emission spectroscopy is a simpler technique than laser-induced fluorescence. By using collection optics to direct light emitted from a desired region of the plume to a spectrometer, and measuring the output while scanning the spectrometer across a wide range of wavelengths. Although this can obtain the same type of information as LIF, it is not as accurate, nor is the spatial precision as good. It does, however, allow measurement of the

entire emission profile simultaneously. giving more knowledge of the energy level population distribution and degree of dissociation of the plume gasses. As frozen-flow losses are a significant problem in arcjet operation, such measurements may be desirable. Although this facility was not designed primarily for emission spectroscopy, the capability does exist.

One novel application for emission spectroscopy was developed during the summer. In the description of the power supply earlier, it was noted that the high-frequency switching power supplies used for operational arcjets produce a significant current ripple. By phase-locking the detection system to the current ripple, we were able to determine that this current ripple is reflected by a ripple in the emission from the plume, with the magnitude of the emission ripple being larger than that of the current ripple. This observation may lead to useful insights into the physics of arcjet operation.

In addition, by measuring the phase shift of the emission ripple at various points in the plume, it was determined that the emission ripple was being convected downstream at the plume velocity. This suggests a simple technique for measuring the velocity of an arcjet plume on an operational or flight-test arcjet, where the weight and complexity of LIF hardware would be prohibitive. By using a simple set of optics and detectors, the time-resolved emission can be measured at many points across the plume, and the resulting data Abel-inverted to give density vs. time at any desired point in the plume. By comparing the phase of the emission ripple at various points, a time-of-flight measurement is obtained, which can be used to construct a crude velocity profile of the entire plume.

RESULTS

All of the necessary elements of the facility are in place and have been tested. An arcjet has been operated in the test chamber, verifying vacuum integrity, cooling, propellant feed, and electric power systems. The laser system is operational and the optical paths have been aligned. The ability of the computer to control all equipment used in the experiment has been confirmed. Finally, authorization to operate the system has been obtained from the necessary agencies. At this point, the facility is ready to begin LIF tests on any available arcjet.

Some emission spectroscopy work has already been performed. As described earlier, the effect of power supply current ripple on plume emission characteristics was explored, and a unique new velocity diagnostic was developed. The results of this work were presented in a paper titled "Performance Effects of Interaction Between a Low-Power Arcjet and its Power Processing Unit", AIAA-92-3238, presented at the Joint Propulsion Conference in Nashville earlier this summer.

CONCLUSIONS

The major accomplishment of the summer was the creation and testing of an arcjet optical diagnostic facility at the Phillips laboratory. This facility includes all equipment necessary to perform LIF and emission spectroscopy experiments on any desired experiment. It is capable of fully characterizing the plume flow field, measuring velocity, temperature, and excited state populations with high accuracy and spatial resolution. Such measurements can be made quite rapidly using an automated control and data-gathering system, allowing easy verification of new arcjet designs and theoretical models. Also, emission spectroscopic techniques were used to examine the effect of power supply current ripple on arcjet operation, and to develop an extremely simple method of measuring the arcjet plume velocity.

ION-MOLECULE REACTIONS AT HIGH TEMPERATURES

Melani Menendez-Barreto
Graduate Student
Department of Physics
University of Puerto Rico at Mayaguez
Mayaguez, PR 00680

Jeffrey F. Friedman
Assistant Professor
Department of Physics
University of Puerto Rico at Mayaguez
Mayaguez, PR 00680

Thomas M. Miller
Professor
Department of Physics and Astronomy
University of Oklahoma
Norman, OK 73019

Final Report for:
Summer Research Program
Phillips Laboratory, Geophysics Directorate

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, DC

August 1992

ION-MOLECULE REACTIONS AT HIGH TEMPERATURES

Melani Menéndez-Barreto
Graduate Student
Department of Physics
University of Puerto Rico at Mayaguez

Jeffrey F. Friedman
Assistant Professor
Department of Physics
University of Puerto Rico at Mayaguez

Thomas M. Miller
Professor
Department of Physics and Astronomy
University of Oklahoma

Abstract

A flowing afterglow apparatus designed for the measurement of ion-molecule reaction rate coefficients at temperatures higher than any previous work was debugged and put into operation during the summer of 1992. Ion-molecule reaction rate coefficients were measured for a variety of systems in the temperature range 300-1200 K: $O^- + H_2, D_2, N_2, CO, NO, \text{ and } CH_4$; $Ar^+ + H_2, O_2, CO, NO, \text{ and } CH_4$; $O_2^+ + CH_4$; and $Cl^- + CH_3Br \text{ and } CH_3I$.

Environmentally Safe Propellants:
Ionophilic Polymer in Liquid Salt Systems

Daniel Lee Fuller
Professor of Chemistry
Department of Chemistry and Physics

Nicholls State University
215 Beauregard Hall
Thibodaux, LA 70310

Final Report for:
Summer Research Program
Phillips Laboratory: Edwards Air Force Base

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington D. C.

September 1992

Environmentally Safe Solution Propellants:
Ionophilic Polymer and Liquid Salt Binder Systems

Daniel Lee Fuller
Professor of Chemistry
Department of Chemistry and Physics

Abstract

Energetic polymeric salts as well as mixtures of polymeric salts and liquid HAN have been investigated for use in environmentally safe solid propellants. HAN is a liquid solution of 95% hydroxylammonium nitrate, 5% ammonium nitrate, and a stabilizer. In this study, polymeric salts have been prepared. Poly(allyl amine) and poly(ethyleneimine) served as the cation source. The anions were nitrates, NO_3^- ; nitroformates, $\text{C}(\text{NO}_2)_3^-$; and dinitramides, $\text{N}(\text{NO}_2)_2^-$. Characterization of these compounds is reported with respect to thermal properties, impact, friction sensitivity, as well as their miscibility, compatibility, and gelatination in HAN.

Binary and ternary mixtures of HAN with selected polymers, prepolymers, plasticizers, and monomers have been prepared. The miscibility, compatibility, and gelatination of these mixtures are presented in the report.

Free radical polymerization of monomers in HAN has been undertaken. Five monomers have been studied. Of these five, acrylamide appears to be the most likely candidate for further consideration as a propellant binder in conjunction with gelatination of HAN to achieve the required mechanical strength needed for a solid rocket propellant.

Environmentally Safe Solution Propellants: Ionophilic Polymer and Liquid Salt Binder Systems

Daniel Lee Fuller

Introduction

Formulation of environmentally safe energetic solid rocket propellants is the goal of rocket scientists. These new environmentally safe propellants must retain or exceed the specific impulse of the present HTPB/isocyanate cured ammonium perchlorate solid propellants. The conventional propellants, utilizing the technologies of the 1950's, had an acid plume of 16-20% HCl. In the 1980's, the scavenged propellants had reduced the acid plume to 1-15% HCl. The present generation of HTPB solid rocket propellants still use ammonium perchlorate hence can not be classified environmentally safe with respect to HCl and they have limited operational safety. The ultimate goal of the formulator is to produce an environmentally safe rocket propellant producing 0% HCl. The goal is to obtain a propellant possessing low explosive hazard for improved operational safety but retaining the high specific impulse critical for rocket propellants.

Although HTPB based propellants have acceptable mechanical properties, they are obtained with great processing difficulty. The processing variables are numerous and, as a result, control of product quality is difficult. The formulators's goal is to simplify the processing required in the preparation of solid propellants. Thus there is a real need to examine new and relatively unknown systems for their use as new solid propellants.

Aerojet Solid Propulsion Company, Sacramento, CA has recently developed a unique propellant system. The oxidizer is a liquid salt. Adding the fuel, a polymer, to the liquid oxidizer produces a solid gel. Curing of the propellant results not from a chemical reaction but from swelling of the polymer by the liquid oxidizer. Clean up does not require ozone depleting chlorinated hydrocarbons as is necessary with HTPB based propellants. Rather, HAN solid gels requires only deionized water for clean up.

Screening tests are needed to identify neutral polymers and polymer salts that might provide adequate solidification through swelling of the liquid oxidizer. It is necessary to identify compounds that are miscible and compatible with liquid oxidizing salts. It is the purpose of this study to identify compounds that will improve the energetics, thermal stability, chemical compatibility, and safety for future solid solution propellant propulsion needs. The current study includes the synthesis of energetic selected ionophilic polymeric compounds, determination of their miscibility and compatibility in HAN, and screening them for impact and friction sensitivity.

Experimental

PAAm, poly(allyl amine), in the free base form was prepared from the hydrochloride salt by neutralization with an excess of aqueous NaOH followed by exhaustive dialysis against deionized water. The PAAm free base was concentrated by vacuum stripping of the solvent.

PEI, poly(ethylene imine), was obtained in the free base form from Dow Chemical.

HAN was obtained from Aerojet Propulsion Division and was used with no further treatment.

PAAMNO₃ was prepared by using a slight excess of nitric acid to an aqueous solution of the PAAM free base described above. The solvent was then removed by evaporation of the solvent in a vacuum oven at temperatures of either 25C or 40C.

PAAMC(NO₂)₃ was prepared following the procedure described for PAAMNO₃ except using 90% methanol and 10% water as the solvent. The nitroformate salts were only very slightly soluble in water.

PAAMC(NO₂)₂ was prepared by a metathesis reaction. Aqueous ammonium dinatramide was added a drop at a time to the aqueous PAAM free base. After stirring overnight, the reaction mixture was placed in a vacuum oven at room temperature to remove the ammonia and water.

Polymerization of acrylamide and copolymerization of acrylamide and N,N'-methylene bisacrylamide was conducted using HAN as solvent and (NH₄)₂S₂O₈ as the initiator. The system was degassed three times using a vacuum system and Liquid N₂. The reaction was carried out at 20°C under a N₂ atmosphere and allowed to react overnight.

The experimental investigation of the thermochemical analysis was conducted by differential scanning calorimetry (DSC) and differential thermal analysis (DTA). Infrared spectral analysis (FTIR) was used to compare the functional groups of the

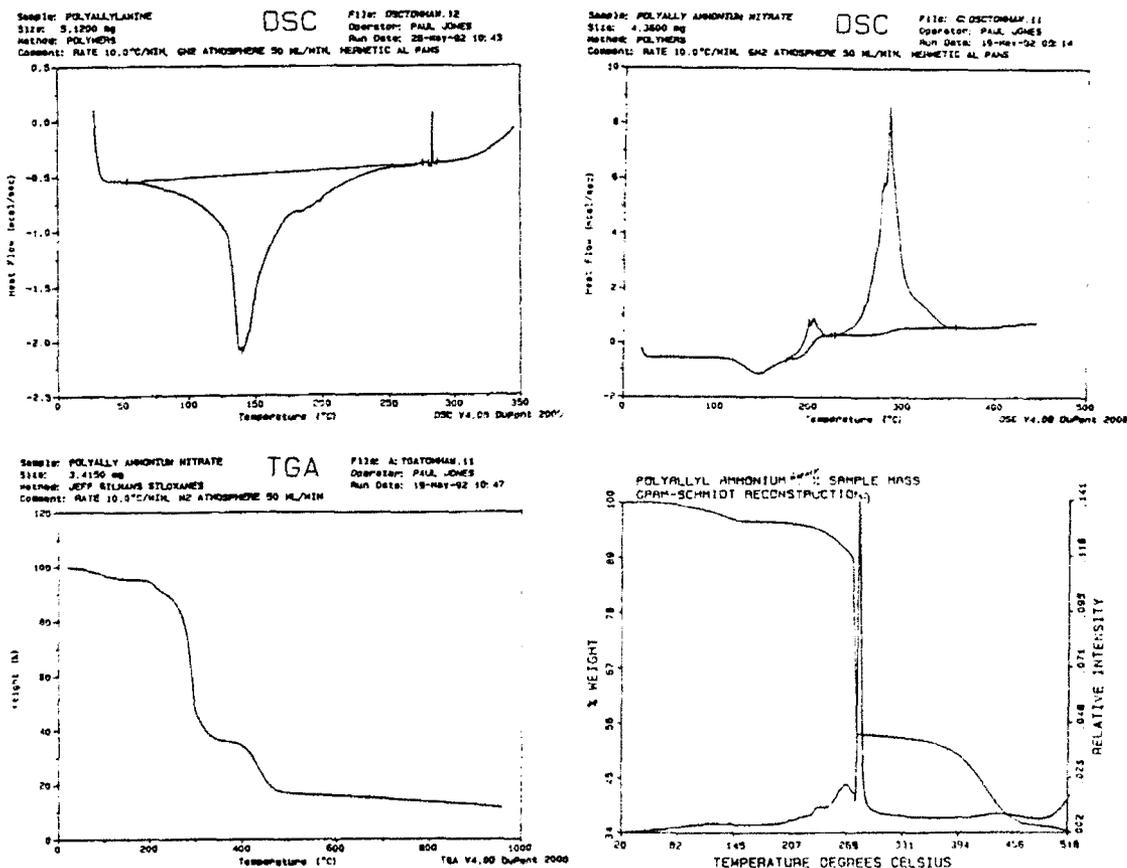
various synthesized polysalts and to determine the structure of the products produced from the thermal degradation of the PAAmNO₃.

Results

1. Energetics of Polymer Salts

Typical thermochemical analysis curves of poly(allyl amine) and poly(allylammonium nitrate) are shown in Figure 1. The poly(allyl amine) shows a slight exothermic, 0.25 cal/g, process at 283°C. The DSC of PAAm/HNO₃(1/1) shown has a small exotherm at 200°C and a significant exotherm (143 cal/g) at 286°C. The thermal decomposition process consists of a two stage weight loss process. The first stage weight loss of the DTA

Figure 1. Typical DSC, DTA, and TG curves of PAAM and PAAmNO₃.



corresponds to the exothermic peak observed in the DSC. The weight loss occurs rapidly between 275 and 293°C accompanied by approximately a 60% weight loss. The weight loss during the second stage process occurs between 408 and 436°C without liberation of heat. The total weight loss for both stages accounts for approximately 80%.

In order to clarify products from decomposition of the ionophilic polymers the decomposition products formed upon heating were examined with the use of FTIR. The gases produced during thermogravimetric analysis of poly(allylammonium nitrate), (1/1) were studied using FTIR. The gases identified in each peak on the thermogram of the TG is listed in Table I. Based on peak area, over 99% of the nitrogen in the PAAmNO₃ was converted to ammonia.

Table 1. Gases Identified from the Decomposition of PAAmNO₃ During Thermochemical Analysis.

Peak	Temp. (°C)	Gases
1	≈210	CO ₂ , NH ₃
2	≈240	CO ₂ , NH ₃
3	≈270	CO ₂ , NH ₃
4	≈300	CO ₂ , NH ₃ , H ₂ O, N ₂ O, HC

Table II illustrates the decomposition temperatures determined by DSC and the heat involved upon decomposition of numerous nitrates, nitroformates, and dinitramides of poly(allyl amine) and nitrates of polyethyleneimine. For the polymeric salt systems studied, the decomposition temperatures did not show any consistent trends with respect to the degree of neutralization of

the amine free base. Further work to refine the synthesis of the polymer salts, purification of the polymer salt, and particle size of the final product will improve the accuracy of the decomposition temperature measurements.

Table II. Decomposition Temperatures and Exothermicity of Ionomers of PAAm and PEI.

Polymer/ Acid	Mole Ratio (Polymer/Acid)	Decomposition Temperature(°C)	Heat (Cal/Gram)
PAAm /HNO ₃	1 / 1	283	143
	1 / 0.75	278	na
	1 / 0.50	295	na
	1 / 0.25	322	na
PAAm /HC(NO ₂) ₃	1 / 1	145	675
	1 / 0.75	145	464
	1 / 0.50	155	447
	1 / 0.25	175	404
PAAm /NH ₄ N(NO ₂) ₂	1 / 1	199	333
	1 / 0.75	201	365
	1 / 0.50	212	291
	1 / 0.25	215	172
PEI /HNO ₃	1 / 1	174	(a)
	1 / 0.75	302	na
	1 / 0.50	(b)	na
	1 / 0.25	(b)	na

na: not available

(a): During decomposition, top kept blowing off

(b): Products were viscous liquids.

As can be seen in Table III, regardless of the extent of nitration, both the PAAmNO₃'s and the PEINO₃'s proved to be insensitive to impact. These conclusions are based on obtaining negative no firings from a 2.5 Kg mass falling a height of 30 cm or greater. The nitroformates proved to be the most sensitive poly salts studied having a sensitivity number, E₅₀, between 25-30 Kg-cm. These values indicate the level at which one-half of the samples tested would be expected to explode. As indicated in Table III, the impact sensitivities of all PAAmDN's

and one-half of the PAAmNF's still need to be determined. It should be expected that the dinitramides might be more impact sensitive than those of either the nitrates or nitroformates since dinitramides are inherently more energetic. As indicated in Table III, the nitrates and nitroformates of PAAm and the nitrates of PEI are insensitive. The friction sensitivity, E_{50} , was 8.4 for both samples of $PEINO_3$. For the nitrates and nitroformates of PAAm, E values of 16.8, 25.2, 32.8, 37.8 Kg all led to negative results.

Table III. Impact and Friction Sensitivity of Polymer Salts.

Polymer / Acid	Mole Ratio Poly./Acid	Impact Sensitivity	Friction Sensitivity
PAAm / HNO_3	1 / 1	Insensitive	Insensitive
	1 / 0.75	Insensitive	Insensitive
	1 / 0.50	Insensitive	Insensitive
	1 / 0.25	Insensitive	Insensitive
PAAm / $HC(NO_2)_3$	1 / 1	Sensitive	Sensitive
	1 / 0.75	Sensitive	Sensitive
	1 / 0.50	Sensitive	Sensitive
	1 / 0.25	Sensitive	Sensitive
PEI / HNO_3	1 / 1	Insensitive	Insensitive
	1 / 0.75	Insensitive	Insensitive

2. Compatibility and Miscibility of HAN Mixtures

Preliminary screening tests were performed for compatibility and miscibility of liquid oxidizer, HAN, and compounds of interest in solid rocket propellant research. The concentration of the additive to HAN was approximately 10% by weight of the total mixture. Assessment of the miscibility, compatibility, and gelatination was qualitative. The evaluation was based principal-

ly on the appearance of the resulting mixture after one hour, twenty-four hours, and one week. The substances studied and the evaluations are given in Table IV. The substances that were miscible in HAN were characteristically those that would be classified as having moderate solubility parameters and capable of either moderate or strong hydrogen bonding. Those polymers or prepolymers that possessed a pendent group incapable of participating in hydrogen bonding were not miscible in HAN.

Table IV. Miscibility, Compatibility, and Gelatination of HAN with Polymeric Compounds and Mixtures of Polymers.

Substance	Miscibility	Compatibility	Gel
<i>Prepolymers</i>			
R45HT(a)	No	Yes	No
GAP 1000-9-2(b)	No	Yes	No
PEG 300(c)	Yes	Yes	No
PEG 600	Yes	Yes	Yes
PEG 1000	Yes	Yes	Yes
PEG 2000	Yes	Yes	Yes
PEG 3000	Yes	Yes	Yes
PEG 3400	Yes	Yes	Yes
Jeffamine 4000(d)	No	Yes	No
Tone P-300(e)	No	Yes	No
PPG 1000(f)	No	Yes	No
PPG 2000	No	Yes	No
PPG 3000	No	Yes	No
PPG 4000	No	Yes	No
Poly(tetramethylene ether glycol) MW 2900	No	Yes	No
<i>Polymers</i>			
Poly(acrylamide)	Yes	Yes	Yes
Poly(vinylalcohol)	Yes	Yes	Yes
Poly(allyl amine)	Yes	Yes	Yes
Polybutadiene (triblock)	No	Yes	No
Poly(ethyleneimine)	Yes	No	Yes
Poly(acrylamide-95- acrylic acid-5- sodium salt)	Yes	No	No

Substance	Miscibility	Compatibility	Gel
<i>Polymer Salts</i>			
PAAm/HNO ₃			
1/1	Yes	No	No
1/0.75	Yes	Yes	No
1/0.50	Yes	Yes	No
1/0.25	Yes	Yes	No
PAAm/HC(NO ₂) ₃			
1/1	Yes	Yes	No
1/0.75	Yes	Yes	No
1/0.50	Yes	Yes	No
1/0.25	Yes	Yes	No
PAAm/HN(NO ₂) ₂			
1/1	Yes	No	No
1/0.75	Yes	Yes	No
1/0.50	Yes	Yes	No
1/0.25	Yes	Yes	No
PEI/HNO ₃			
1/1	Yes	Yes	No
1/0.75	Yes	Yes	No
1/0.50	Yes	Yes	No
1/0.25	Yes	Yes	No
<i>Plasticizer</i>			
TEGDN(g)	No	Yes	No
TMETN(h)	No	Yes	No
<i>Isocyanates</i>			
TDI(i)	No	Yes	No
IPDI(j)	NO	Yes	No

(a) hydroxyterminated polybutadiene-45% 1,2-adduct/55% 1,4-adduct.
 (b) poly(glycidyl azide) (c) poly(ethylene glycol). (d) amine terminated poly(propylene glycol). (e) Hydroxy terminated polycaprolactone. (f) poly(propylene glycol). (g) triethylene glycol dinitrate.
 (h) trimethylol ethane trinitrate. (i) Toluene diisocyanate.
 (j) isophorone diisocyanate.

Of the various prepolymers, isocyanates, and plasticizers examined, all were found to be chemically compatible although not always miscible with the liquid oxidizer HAN. Four polymers were

found to be chemically incompatible. Three of these polymers were polymer salts.

Most of the various polymeric salts tested with HAN did not result in the formation of a gel. The resulting solutions were observably more viscous than the pure HAN. For a polymer salt, as

Table V. Compatibility and Gelation of Ternary Mixtures of Liquid HAN. [$\approx 90\%$ HAN and $\approx 10\%$ Solutes by Weight]

Weight Ratio		Description
Polymer Salt/Neutral Polymer		of Mixture
<i>PAAmNO₃</i> and <i>PAAm</i>		
3	/ 1	liq. HAN/Swollen Salt
2	/ 2	liq. HAN/Swollen Salt
1	/ 3	Solution
<i>PAAmC(NO₂)₃</i> and <i>PAAm</i>		
3	/ 1	Decomposition of Liq./Swollen Salt
2	/ 2	Decomposition of the Mixture
1	/ 3	Decomposition of the Mixture
<i>PAAmN(NO₂)₂</i> and <i>PAAm</i>		
3	/ 1	liq. HAN/Swollen Salt
2	/ 2	liq. HAN/Swollen Salt
1	/ 3	liq. HAN/Swollen Salt
<i>PAAmNO₃</i> and <i>PAA</i>		
3	/ 1	Very Soft Gel
2	/ 2	Soft Gel
1	/ 3	Firm Gel
<i>PAAmC(NO₂)₃</i> and <i>PAA</i>		
3	/ 1	Soft Gel
2	/ 2	Firm Gel
1	/ 3	Very Firm Gel
<i>PAAmN(NO₂)₂</i> and <i>PAA</i>		
3	/ 1	Firm Gel
2	/ 2	Very Firm Gel
1	/ 3	Gel with Elastic Properties
<i>PEINO₃</i> and <i>PAA</i>		
3	/ 1	Very Soft Gel
2	/ 2	Soft Gel
1	/ 3	Firm Gel

the percent of salt to free base decreased in the compound, the viscosity steadily increased. Poly(acrylamide), PAA, is superior to poly(allyl amine) in the ability to gel liquid HAN. PAA leads to the formation of gels in both binary and ternary mixtures with HAN. Table V shows the firmness of the gels formed with PAA increase with decreasing polymer salt. From the information given in Table V, it becomes apparent that mixtures of PAAM does not form gels in ternary mixture with ionophilic polymers and HAN.

3. Polymerization Reactions In HAN

Since a mixture of HAN and polyacrylamide produces a gel, could polymerization of an acrylic monomer in HAN yield a more uniform gel? Polymerization would allow the possibility of introducing cross-linking and allow the formulator a means of improving the mechanical properties of the propellant. Seven acrylic monomers were investigated. The seven acrylic monomers investigated and the results of the polymerization reaction are summarized in Table VI. The polymerization solutions were prepared by combining the acrylic monomer with HAN and ammonium peroxydisulfate; the mixture was allowed to react overnight at room temperature. The procedure used was to mix the HAN, monomer, and initiator. The mixture was then degassed by freezing and thawing three times using liquid N_2 as the coolant. Although the monomer was soluble in HAN, the initiator was only partially soluble. The mixture then was set aside for overnight. By the following day, the initiator no longer remained as a white solid. A 1 to 53 by weight ratio of N,N'-methylenebisacrylamide to acrylamide in HAN

containing the initiator yielded a gel with significantly superior improved properties over that of the acrylamide only. Presently, no explanation for why only two of the acrylic monomers led to polymerization.

Table VI. Acrylic Monomers Tested in this Study.

Monomer	% Mass	Miscible	Change in State
N,N-Dimethyl Acrylamide	9.9	Yes	None
N-Methyl Acrylamide	9.9	Yes	None
Acrylamide	10.2	Yes	Solid
N-Methyl Methacrylamide	10.1	Yes	None
N,N'-Methylene Bisacrylamide	9.5	Yes	Solid
N-Vinyl-N-Methyl Acetamide	9.9	Yes	None
4-Vinyl-Pyridine	9.8	Yes	None

Comonomers	% Mass	Miscible	Change in State
Acrylamide	0.17	Yes	Elastic Solid
<u>N,N'-Methylene Bisacrylamide</u>	<u>9.67</u>		

References

- (1) Lewis, E.A.; Barkely, T.J.; Feams, R.R.; Hansen, L.D.; St.Pierre, T., *Macromolecules*, 1984, 17, 2874.
- (2) Lovett, J.R., Rassannate, A.S., U.S. Pat. 3,401,156 [C.A. 69 87714g].

AN INVESTIGATION OF THE FEASIBILITY OF VARIOUS
ENERGETIC SALT COMBINATIONS FOR USE IN SOLUTION PROPELLANTS

Vincent P. Giannamore
Assistant Professor
Department of Chemistry

Xavier University of Louisiana
7325 Palmetto Street
New Orleans, LA 70125

Final Report for:
AFOSR Summer Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, DC

September 1992

AN INVESTIGATION OF THE FEASIBILITY OF VARIOUS
ENERGETIC SALT COMBINATIONS FOR USE IN SOLUTION PROPELLANTS

Vincent P. Giannamore
Assistant Professor
Department of Chemistry
Xavier University of Louisiana

Abstract

Mixtures of energetic salts were examined for their feasibility of use in solution propellants. The principle goal of this initial study was to discover mixtures which were liquid near 298 K. Six different combinations of salts were investigated. These included ammonium dinitramide/ammonium nitrate, ammonium dinitramide/hydrazinium nitrate, ammonium dinitramide/guanidinium nitroformate, ammonium dinitramide/hydrazinium nitroformate, guanidinium nitroformate/ammonium nitrate, and guanidinium nitroformate/hydrazinium nitroformate. A total of thirty-three different compositions of these mixtures were examined. None of the mixtures studied were entirely satisfactory. However, some appear to be worth further study. Mixtures of ammonium dinitramide and ammonium nitrate containing 75 mole percent and 90 mole percent ammonium dinitramide appear to be the most promising.

AN INVESTIGATION OF THE FEASIBILITY OF VARIOUS ENERGETIC SALT COMBINATIONS FOR USE IN SOLUTION PROPELLANTS

Vincent P. Giannamore

INTRODUCTION

There are several potential sources of environmental hazards associated with rocket propellants. These include hazardous waste from the manufacturing process, toxic exhaust, and disposal. All phases of a propellant's life cycle; including initial research and development, motor production, motor testing, and disposal and demilitarization; are involved. Issues related to environmental propellant have been discussed and assessed recently by Hawkins and Wilkerson [1].

One of the areas of concern is the presence of hydrochloric acid in solid rocket exhaust. Political concerns and public perception are major factors in this area. Evaluation of the effects of these rocket emissions by technical professionals in the environmental and propulsion communities indicated that the environmental effects were relatively small. Nevertheless, it was concluded that reduction of undesirable environmental effects should be pursued in future rocket propulsion systems. This approach is consistent with EPA priorities which stress that reduction of hazardous waste generation is preferable to treatment and disposal.

Among a number of new technologies being investigated for feasibility of use in future systems are liquid eutectic salt solutions of energetic salts. Solution propellant research at Edwards AFB this summer was divided into three categories: oxidizer development, polymer development, and theoretical performance. This paper reports on the oxidizer development category. Various combinations of energetic salts were investigated. Since this Energetic Salt Solution Technology (ESST) is still new, research is in its initial stages. The principle goal was to discover oxidizer salt combinations which are liquid near 298 K (25 °C).

EXPERIMENTAL

Room temperature molten salt solutions have been investigated by groups at several locations including the Air Force's Seiler Laboratory. However, to our knowledge, none of these investigations have involved energetic materials. Preparation of the mixtures is generally straightforward. The salts are mixed in the proper proportions and stirred thoroughly until the mixture liquefies.

This technique was not practical in the present study, however. In the first place, it was not known what ratio, if any, of the two salts would produce a low temperature eutectic mixture. Thus, one could not know whether continued stirring would produce practical results. A more serious impediment to application of this method in the present case was the sensitivity of energetic materials to impact and friction. Vigorous stirring of the mixtures was

not advisable for safety reasons. Therefore, another method of producing well-mixed combinations of the various salts had to be employed.

The following method was employed. Both solid salts were placed in a small container, usually a test tube. The total mass was approximately 0.1 g. The small scale was employed because nothing was known about the sensitivity of the mixtures to various stresses. The mixture was heated in an oil bath behind a shield. Heat was applied until melting occurred and the mixture was all liquid or until it did not appear prudent to continue heating. (See discussion.) In this way it was hoped that thorough mixing of the two compounds would take place. The mixtures were then observed as they cooled to room temperature.

Melting points were obtained on a Thomas-Hoover Melting Point Apparatus and temperatures are uncorrected. Thermal analysis was performed on a DuPont Differential Scanning Calorimeter. Three methods were employed. In all methods the sample was placed in hermetically sealed aluminum pans and a gaseous nitrogen atmosphere with a flow rate of 50 mL/min was employed. In method 1 samples were heated at a rate of 10 °C/min from ambient temperature to 350 °C. Method 2 was identical except that the sample was cooled with liquid nitrogen prior to heating. Method 3 involved heating the sample to 125 °C, then cooling it to 35 °C, followed by heating to 350 °C. Infrared analyses were performed on a Nicolet PAS-FTIR

RESULTS AND DISCUSSION

The first mixtures examined were composed of ammonium dinitramide (ADN) and ammonium nitrate (AN). The compositions of the mixtures are given in the table below.

Mixture	Mass of ADN	Mass of AN	Mole % ADN
1	0.0147 g	0.0861 g	10
2	0.0344 g	0.0665 g	25
3	0.0531 g	0.0523 g	39
4	0.0607 g	0.0397 g	50
5	0.0821 g	0.0178 g	75
6	0.0831 g	0.0166 g	76
7	0.0936 g	0.0069 g	90
8	0.0932 g	0.0059 g	91

The ADN reportedly contained a small amount of AN from the manufacturing process. The AN used for these mixtures, but not for those reported later, contained a small, but unspecified amount of clay. The masses reported in the table do not take these impurities into account.

Mixture 1 was prepared by heating the mixed salts to 162 °C. Some evolution of gas was noted near that temperature. This mixture did not melt below 125 °C.

Mixture 2 was prepared by heating the mixed salts to 161 °C. The mixture was completely melted by 110 °C when checked in the Thomas-Hoover apparatus. A DSC thermogram showed an endotherm at 94 - 108 °C which was not present in thermograms of either AN or ADN.

Mixture 3 was prepared by heating the mixed salts to 166 °C. The mixture in the test tube appeared to be solid at 166 °C, but, upon cooling, it appeared liquid at 100 °C and solid again by 50 °C. A melting point check showed it to be mostly liquid by 60 °C, although some solid persisted to much higher temperatures. The DSC showed an endotherm at 49 - 54 °C. This occurs at a temperature similar to an endotherm in the DSC of AN. However, the energy of the present transition is greater.

Mixture 4 was prepared by heating the mixed salts to 164 °C. The mixture was entirely melted by 70 °C. DSC showed an endotherm at 61 - 64 °C.

Mixture 5 was prepared by heating the mixed salts to 98 °C. The entire sample appeared to be liquid at that temperature, so it was not heated any further. A melting point check on the Thomas-Hoover apparatus showed that the sample was nearly completely melted by 75 °C but a few particles still remained solid at 120 °C. Recall that the original AN contained a trace of clay. The DSC showed an endotherm at 50 - 67 °C. Oddly, a sample which was heated to 125 °C in a capillary tube did not appear to solidify until it had cooled to 35 °C.

Mixture 6 was prepared by heating the mixed salts to 110 °C. No melting point or DSC thermogram was obtained on this sample. An attempt was made to observe the mixture closely while it cooled. It appeared to be liquid down to 30 °C. At 29 °C close inspection revealed that it had solidified. Because of the way the apparatus was set up, the mixture could well have been a gel before that temperature but still have appeared liquid.

Mixture 7 was prepared by heating the mixed salts to 96 °C, at which point the entire sample appeared to be liquid. As it cooled it appeared to be liquid until 41 °C and solidify rapidly at that temperature. When a sample was checked on the melting point apparatus, its melting point was well above 41 °C. However, when this sample was observed as it cooled, it rapidly solidified at 41 °C. This was checked twice. None of the other mixtures tested in this way (1,2,4, and 5) exhibited this rapid solidification, although mixture 5 was liquid at a lower temperature as reported above. An attempt was made to correlate this behavior with DSC thermograms obtained while the samples were being cooled (method 3). Mixture 7 showed an exotherm which returned to the baseline at 41 °C. However, thermograms of mixtures which did not exhibit the same solidification behavior also showed similar exotherms.

Mixture 8 was prepared by heating the mixed salts to 118 °C. No melting point nor DSC was obtained. Observation of the sample upon cooling showed similar behavior to mixture 6. In this case, however, the sample was examined more closely and appeared to be gel-like at 38 °C. At 29 °C it still appeared the same, but after storage in a desiccator overnight, it was found to be a white solid. The difficulty of making good observations of mixtures 6 and 8 in the apparatus used must be stressed.

The next set of mixtures investigated were composed of ADN and guanidinium nitroformate (GNF). The composition of the mixtures examined is given in the table below.

Mixture	Mass of ADN	Mass of GNF	mole % ADN	mass % ADN
9	0.0106 g	0.0899 g	17	11
10	0.029 g	0.077 g	39	27
11	0.047 g	0.046 g	63	51
12	0.075 g	0.027 g	83	74
13	0.0892 g	0.0097 g	94	90

These mixtures were prepared similarly to mixtures 1 through 8, although temperatures employed did not exceed 130 °C in any of these cases. Generally, the samples were heated to 100 - 110 °C to prepare the mixtures. None of these mixtures appeared to be liquid in the desired range. The lowest melting point was observed for mixture 10, which melted at 64 - 76 °C. DSC thermograms did not show endotherms below 90 °C for any of these mixtures. Interestingly, only the DSC of mixture 9 shows an exotherm in the 127 °C range. DSC thermograms of GNF show large exotherms in this area. This is interesting since the ADN/GNF mixtures appear to have exhibited much more thermal stability than other GNF mixtures (see below). Perhaps the presence of ADN stabilizes the GNF in some way.

The third set of mixtures examined were composed of ADN and hydrazinium nitrate (HN). The HN was prepared by the reaction of hydrazine with ammonium nitrate. The procedure involved slowly adding the ammonium nitrate to hydrazine and then removing ammonia at reduced pressure. (2) It did not seem to work well when 1:1 ratios of reactants were used. The best results were obtained when an excess of ammonium nitrate was used. The product had a satisfactory infrared spectrum but it appeared wet and had a low melting point. More work in this area was required but time did not permit it. The table below lists the compositions of the mixtures prepared.

Mixture	Mass of ADN	Mass of HN	Mole % ADN	Mass % ADN
14	0.009 g	0.090 g	7	9
15	0.025 g	0.073 g	21	26
16	0.051 g	0.054 g	42	49
17	0.075 g	0.028 g	67	73
18	0.089 g	0.010 g	87	90

Mixtures 14 through 18 were prepared by mixing the salts and then heating them. All were heated at the same time to 75 °C. None were liquid at that temperature. Each was then individually heated until it had definitely melted. None appeared to hold promise of being liquid near 298 K. Their DSC thermograms appeared to be composed primarily of the same transitions as the individual components.

The fourth set of mixtures were composed of ADN and hydrazinium nitroformate (HNF). The compositions are given in the table below.

Mixture	Mass of ADN	Mass of HNF	Mole % ADN	Mass % ADN
19	0.012 g	0.092 g	16	12
20	0.024 g	0.076 g	32	24
21	0.050 g	0.052 g	59	49
22	0.078 g	0.026 g	82	75
23	0.089 g	0.009 g	94	91

Mixture 19 was prepared in the usual way. The salts were mixed and heated. At 115 °C bubbling was noted in the sample. It was cooled and its melting behavior observed in the Thomas-Hoover apparatus. It did not melt but appeared to decompose with evolution of gas at 100 - 110 °C. Its DSC shows a complex of overlapping exotherms beginning at approximately 110 °C.

Mixture 20 was prepared by heating the mixed salts to 100 °C. It was cut off at that point because bubbling was noted in the sample even though it had not completely liquefied. When a small sample was observed in the melting point apparatus, it softened at 50 - 60 °C but did not melt. It decomposed similarly to mixture 19.

Mixture 21 was mixed and heated to 120 °C. It was liquid at that point but some bubbling was noted. Its behavior in the melting point apparatus was similar to mixtures 19 and 20. It softened near 50 °C and decomposed at 100 - 110 °C.

Mixture 22 was prepared in the same way as mixture 21. It appeared to soften at 53 °C but did not melt. It was not heated to decomposition.

Mixture 23 was prepared by heating the mixed salts to 131 °C. It did not appear to be thoroughly mixed as some parts appeared yellow while others were more orange. There appeared to be some liquid condensed on the upper regions of the test tube suggesting that a gas had been evolved but no bubbling had been noticed. In the melting point apparatus this mixture seemed to melt fairly sharply at 74 - 81 °C. Some bubbles were noticed as the melting occurred.

The fifth set of mixtures were composed of GNF and AN. The compositions are given in the table below.

CONCLUSIONS

As noted previously, the goal of this research was to discover mixtures of energetic salts which are liquid near 298 K. The study was essentially a preliminary scanning of a number of possibilities to discover mixtures which showed promise for further development.

None of the mixtures studied were definitely liquid at 298 K. The ADN/AN mixtures appeared most worthy of further study. Mixtures 5, 6, 7, and 8 were the most promising. It may be worth studying a number of mixtures with compositions in the same range. The mixtures typically did not melt at very low temperatures but once melted could be cooled to relatively low temperatures before solidifying. It would be preferable if the mixture actually melted near room temperature. Further investigation in this range of compositions may uncover such a mixture. However, it could be practical to use a supercooled liquid mixture such as these appear to be. Investigation of the stability of these mixtures would be required. If the liquid state persisted during processing, these mixtures could be suitable for use. Investigation of the sensitivity of the mixtures to impact and friction is also required.

None of the other mixtures seem worthy of further development. Certainly some of the mixtures containing GNF and/or HNF appear to be too thermally unstable for their use to be feasible even if they could be liquefied at an acceptable temperature. Mixtures containing HN may be worth further investigation if a reliable source of HN were employed. Certainly, this is obtainable, and ADN/HN mixtures should be studied further.

All of the mixtures studied were composed of two energetic salts. Two different types of mixtures should be investigated as well. One type is mixtures of energetic salts with small amounts of nonenergetic materials. For example, AN is known to form a eutectic mixture with sodium nitrate. The energetics of this and other such mixtures needs study. The other type which should be investigated is ternary mixtures.

Finally, the method of preparing the mixtures was unsatisfactory. A better method must be developed.

REFERENCES

1. Hawkins, T.W. and Wilkerson, B.E., Environmental Propellant: Current Issues and Assessment, 1992 JANNAF Propulsion Meeting.
2. Owens, T., personal communication

DEVELOPMENT OF A PROTOTYPE LIDAR SYSTEM
AT THE STARFIRE OPTICAL RANGE

Gary G. Gimmestad
Principal Research Scientist
Georgia Tech Research Institute

Georgia Institute of Technology
Atlanta, GA 30332

Final Report for:
Summer Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, DC

September 1992

DEVELOPMENT OF A PROTOTYPE LIDAR SYSTEM
AT THE STARFIRE OPTICAL RANGE

Gary G. Gimmestad
Principal Research Scientist
Georgia Tech Research Institute
Georgia Institute of Technology

Abstract

A prototype lidar system has been developed and demonstrated at the Starfire Optical Range (SOR). The lidar system uses a copper vapor laser (CVL) as a source. This laser has two output wavelengths, 0.5106 and 0.5782 microns, with roughly equal power. The laser beam is projected through a 1.5 meter telescope and is focussed at 10 km range. The lidar receiver is based on a 14-inch telescope mounted on the side of the 1.5 meter telescope. The receiver has two optical channels and the lidar can be configured either as a two-color system or as long-range and short-range receivers which span an altitude range from the planetary boundary layer through the stratospheric aerosol layer. The receiver components are mounted on an optical breadboard to allow easy changes in configuration. The two-channel data system has 12-bit transient recorders with a maximum digitization rate of 20 MHz and digital memories capable of averaging up to 64 k sweeps. The digitizers were interfaced to a PC AT computer which was used for data acquisition, recording, analysis and display. The system was normally operated with 10 Mhz digitization rate, which yields 15 m range resolution, and 3000-sweep averaging which required 3.6 seconds per measurement. The maximum range of the system was limited to 30 km by the 5 kHz pulse repetition frequency of the CVL. The system was demonstrated by monitoring subvisual cirrus drifting across the zenith sky with 30-second time resolution. Analysis of this data showed that the transmittance of the cirrus, which ranged from 0.9 to 1.0, could be determined with an accuracy of about +/- 0.02.

DEVELOPMENT OF A PROTOTYPE LIDAR SYSTEM
AT THE STARFIRE OPTICAL RANGE

Gary G. Gimmestad

INTRODUCTION

The Starfire Optical Range (SOR) employs high-power lasers to create artificial guide stars in the atmosphere. These laser guide stars are used in conjunction with adaptive optics techniques to provide high resolution images of celestial objects and man-made satellites. The adaptive optics techniques are used to remove the effects of atmospheric turbulence, but in applications where absolute radiometric measurements are of interest (such as stellar photometry), it is also desirable to remove the effect of attenuation of light as it passes through the atmosphere. This can be done by employing the laser remote sensing technique known as LIDAR (Light Detection And Ranging). In early 1992 the author of this report and the lab focal point, Dr. Robert Q. Fugate, developed a plan to build and evaluate a prototype lidar system at SOR. This plan was carried out under the AFOSR Summer Faculty Research Program during an eight-week period beginning on July 27 and ending on September 18, 1992.

A basic lidar system can be thought of as a transmitter, a receiver, and a data subsystem. The guide star lasers at SOR can be used as lidar transmitters, so development of a lidar system required only the addition of a receiver and a data subsystem, and some of the components for these were available at SOR or on loan from Georgia Tech. The remaining components were purchased by SOR or fabricated in the Phillips Laboratory machine shop.

The objectives of the work described here were to develop, operate and evaluate the prototype lidar at SOR. We had also hoped to demonstrate the usefulness of the lidar in stellar photometry.

DISCUSSION OF PROBLEM

The transmittance of the earth's atmosphere along slant paths from space to the ground is highly variable, because of variations in the atmospheric constituents which absorb or scatter light. These include both gasses and particulate matter. Most of the particulates are in the troposphere, and they range from boundary layer aerosols to various types of clouds including high, thin cirrus. There is also a permanent aerosol layer in the stratosphere ranging from about 12 to 24 km altitude, and this layer can significantly affect transmittance after major volcanic events such as the eruption of Mt. Pinatubo in 1991. The most variable gaseous constituent is water vapor, which is strongly absorbing at certain infrared wavelengths.

Fortunately, a lidar system to support observations at SOR does not have to accommodate all atmospheric conditions, because photometric measurements are only done when the sky is fairly clear. The main issue is obscurants such as subvisual cirrus, which can affect measurements while remaining undetected by observers. Cirrus can be highly variable in its spatial distribution and it moves with high-altitude winds, so patches of cirrus tend to drift across the line of site, introducing artifacts into photometric measurement data. Changes in the concentration of boundary layer aerosols can also affect atmospheric transmittance, although perhaps on a longer time scale because the boundary layer tends to more uniform than cirrus. As mentioned earlier, water vapor is both strongly absorbing in the infrared and highly variable, although astronomical measurements are generally done in the window regions between absorption features in order to minimize the effects of water vapor.

The Starfire Optical Range is equipped with a 1.5 meter telescope with a Coude path. A laser beam from a copper vapor

laser (CVL) is routinely propagated out through the telescope to form an artificial guide star. A 14-inch Celestron telescope is mounted on the side of the 1.5 meter telescope and is boresighted to it. The CVL guide star beam was selected to be the lidar transmitter and the Celestron was selected to be the basis for the lidar receiver. This combination has the advantages that the lidar will always be aimed in the direction the telescope is pointing; that it does not require additional optics in the Coude path; and that operation of the lidar does not interfere with any other system. It has the disadvantage that many of the lidar design parameters are fixed by these choices, so the design is somewhat constrained. For example, the pulse repetition frequency of the CVL is 5 kHz, so the maximum range of the lidar is 30 km, because this is the distance a pulse of light travels (out and back) during the 200 microsecond interval between pulses. Another unusual feature of the CVL is that it has two output wavelengths with roughly equal power in each one. Specifications of the CVL guide star and of the Celestron telescope are listed in Tables 1 and 2.

Table 1. Specifications of CVL Guide Star

Wavelengths:	0.5106 and 0.5782 microns
Power transmitted:	70 Watts (typical)
Pulse repetition rate:	5 kHz
Energy per pulse:	7 mJ each wavelength (typical)
Pulse width:	50 ns
Beam geometry:	1.5 m diameter leaving the telescope, focussed at 10 km range, beam diameter 10-20 cm at focus.

Table 2. Specifications of Celestron Telescope

Model:	C14
Optical design:	Schmidt-Cassegrainian catadioptric
Primary mirror:	14.25 inch dia., 60 inch radius, spherical
Secondary mirror:	3.5 inch dia., 17.6 inch radius, aspheric
Central obstruction:	4.5 inches
Focal ratio:	f/11
Focal length:	154 inches
Image scale:	0.256 mrad/mm

METHODOLOGY

Several options were considered for the prototype lidar, including water vapor profiling by either the DIAL technique (Differential Absorption Lidar) or by Raman lidar. Neither of these proved feasible given the laser specifications in Table 1. An investigation of the water vapor spectrum using the HITRAN PC software showed that neither of the CVL lines is on a water vapor absorption feature. The Raman technique relies on inelastic scattering by water vapor and nitrogen, which causes characteristic frequency shifts in the scattered light. Unfortunately, the nitrogen-shifted 0.5106 micron line is very near the 0.5782 micron laser line and in any event the Raman scattering cross sections are very small and the energy per pulse from the CVL is also small, so such a system would have a very limited altitude range.

The prototype lidar was therefore designed to monitor atmospheric particulates. As mentioned earlier, both boundary layer aerosols (at the bottom of the troposphere) and cirrus

clouds (near the top of the troposphere) can cause changes in atmospheric transmittance during the course of an observation, so the lidar was designed to monitor both of these simultaneously. This necessitated two receiver channels in order to have sufficient dynamic range (and also to avoid defocus problems with the Celestron telescope). The wavelength dependence of backscatter from particulates was also thought to be sufficiently interesting to justify a two-color receiver option, so this was included as well. The receiver was constructed on optical breadboard so that its configuration could be easily changed.

The two receiver configurations are shown schematically in Fig. 1. The two-color configuration is shown at the left. The beam divider (BD) is a dichroic which transmits the yellow light (0.5782 microns) to PMT1 and reflects the green light (0.5106 microns) to PMT2. Each of the PMTs is equipped with a narrow-band

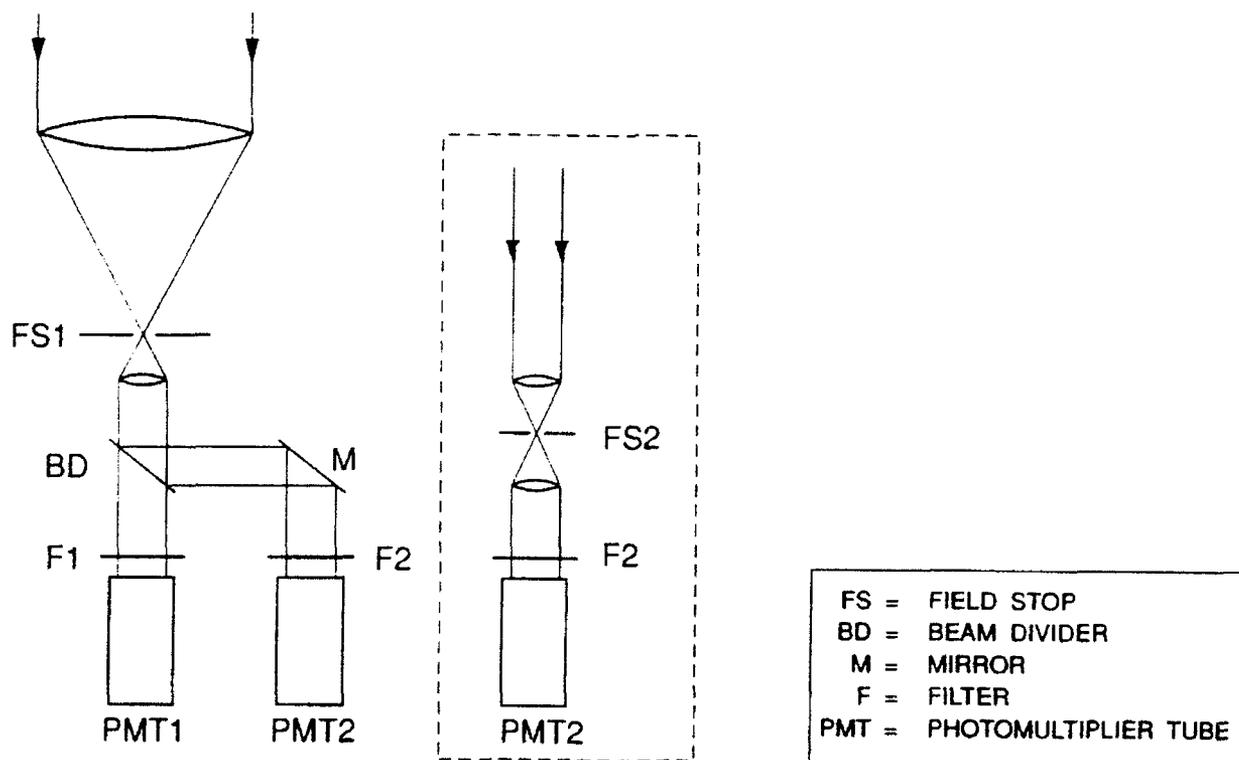


Figure 1. Optical Receiver Schematic

interference filter to reject any residual light from the other channel, and the field stop (FS1) determines the field of view of both channels. In the long range/short range configuration, PMT2 is moved to a receiver based on a 2-inch diameter objective lens, as shown in the dashed box in Fig. 1. Its (wider) field of view is determined by a field stop FS2. Both PMTs had 0.5106 micron filters in this configuration.

The two-channel data subsystem, shown schematically in Fig. 2, was provided on loan by Georgia Tech. This subsystem included software for acquiring, storing, and analyzing data.

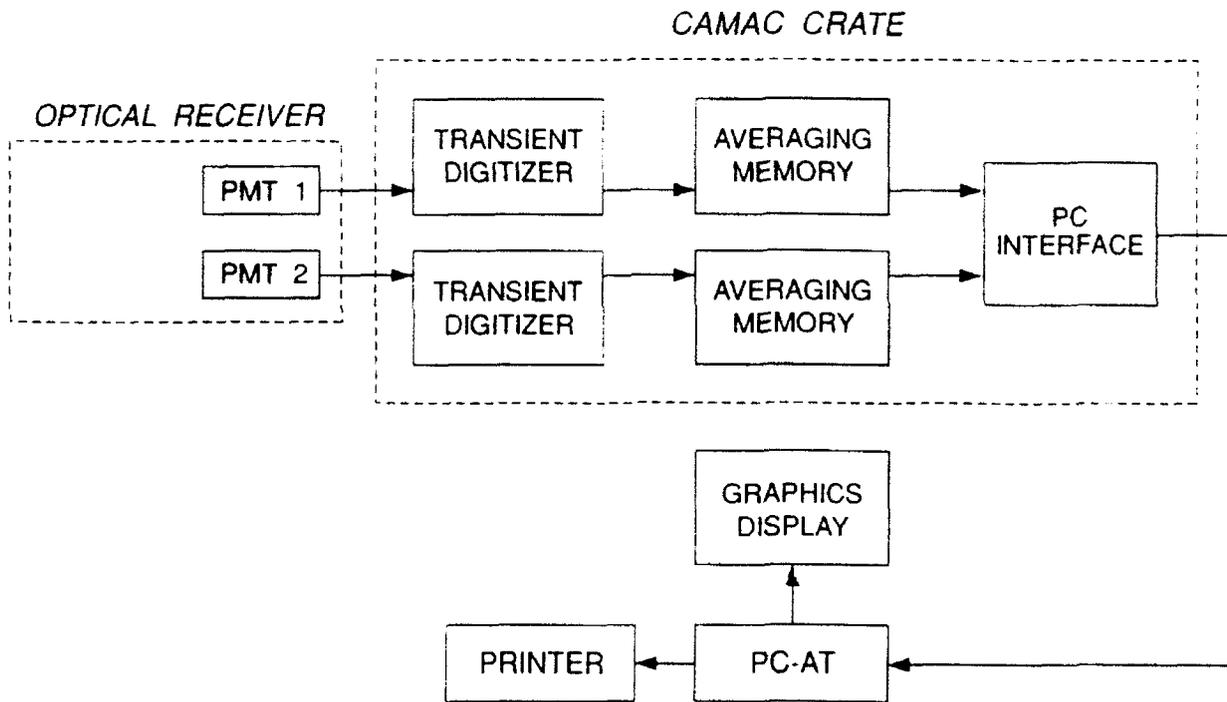


Figure 2. Data Subsystem Block Diagram

RESULTS

Many of the lidar system design parameters were fixed by the decisions to use the CVL guide star as the transmitter, to use the 14-inch Celestron as the receiver telescope, and to use

available components as much as possible. The one key parameter which remained, receiver field of view, is related to the system's dynamic range. Designing a system with sufficient dynamic range is a standard problem in lidar, for two reasons: for a given backscatter coefficient, the amount of light collected by the receiver telescope decreases as inverse range squared; and the Rayleigh backscatter from air molecules is proportional to air density, which decreases exponentially with altitude, with a scale height of about 8 km. These two factors make the signal from nearby scatterers much larger than the signal from distant scatterers. In order to maximize sensitivity at long ranges, it is necessary to somehow discriminate against short ranges, so that short-range signals do not saturate the detector or electronics. One way to do this is to restrict the receiver field of view. Restricting the field of view has the added virtue of decreasing the amount of sky background radiation which reaches the detector. The distance at which the transmitted laser beam is fully within the receiver field of view is determined by the laser beam geometry, the separation of the transmitter and receiver optic axes, and the receiver field of view. The first two of these three parameters were fixed, so only the field of view was an adjustable parameter. The specifications of the receiver and data subsystem are listed in Tables 3 and 4, respectively.

The prototype SOR lidar system was operated on 13 occasions during the period August 13 to September 18, 92. The first three times were for system check-out, and during these, the receiver was mounted on a tripod and placed outdoors near the CVL laboratory, aimed at the zenith. The CVL beam was brought out through a hole in the wall to a turning mirror on an adjustable mount. The distance between the laser beam axis and the receiver axis was set so that the crossover distance was the same as it would be later when the 14-inch receiver telescope was mounted on the 1.5 meter telescope.

Table 3. Specifications of Lidar Receiver

	Long Range	Short Range
Aperture	14 inches	2 inches
Field of View	0.8 mr	11 mr
PMT	S-20	S-20
Optical filter		
center wavelength	0.511, 0.578	0.511 microns
bandpass	1 nm	1 nm
peak transmittance	60 %	60 %

Table 4. Specifications of Data Subsystem

Digitization rate:	20 MHz (max.)
Resolution:	12 bits
Averaging:	64 k sweeps (max.)
Storage & display	PC AT

Typical results from the long-range receiver on August 18, 1992 are shown in the next three figures. The raw data is shown in Fig. 3. This data is the average of 3000 laser pulses, which required 3.6 seconds to acquire (the data system only uses each sixth laser pulse due to internal timing requirements). The digitization rate was 10 MHz, corresponding to 15 meter range bins. The signal increases with altitude at first as the laser beam enters the receiver field of view, then it falls rapidly due to the inverse range squared dependence and the decrease of air

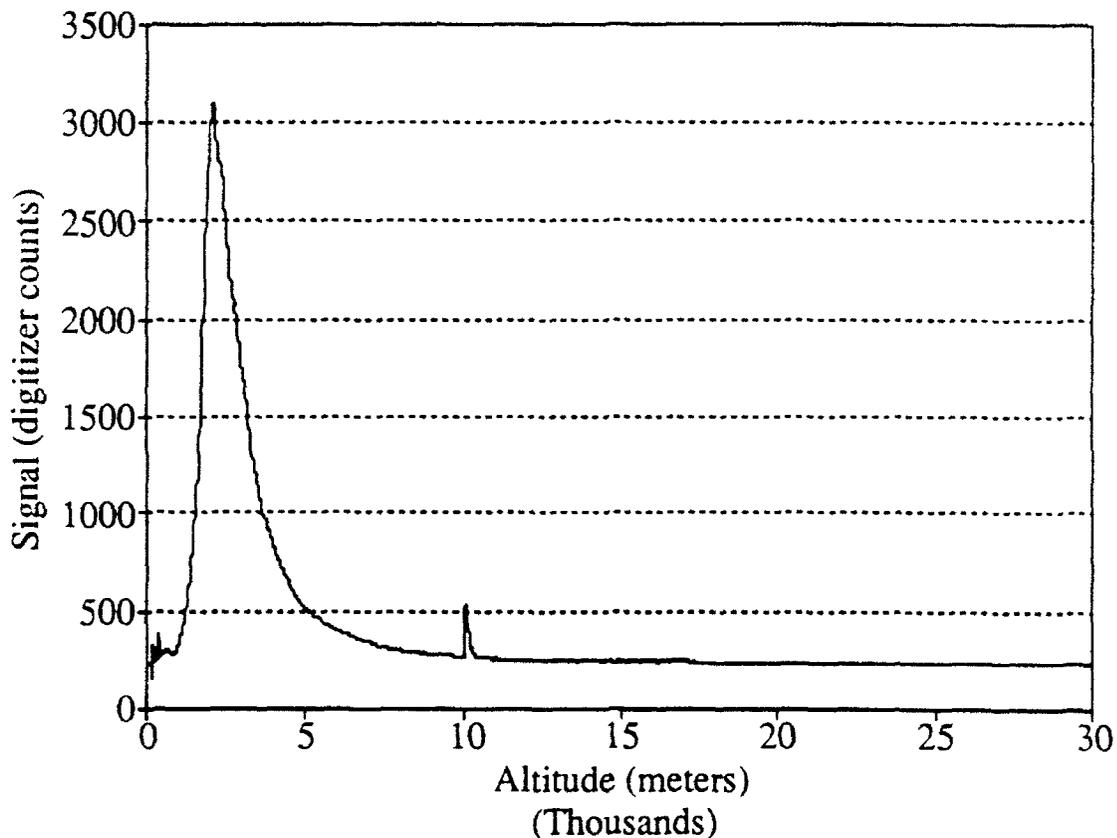


Figure 3. Example of raw data from August 18, 1992.

density with altitude. A signal from a cloud is seen at 10 km, then the signal asymptotically approaches a constant value which is determined by arbitrary offsets in the amplifiers.

The data is easier to interpret if the range dependence is removed. This is done by subtracting the arbitrary offset from all range bins and then multiplying them by range squared. The result, referred to here as "normalized data", is shown in Fig. 4. If there were no particulates in the atmosphere, this curve would be due to Rayleigh scattering by air molecules only, and so it would show only a smooth, exponential decrease with altitude, due to decreasing air density. At altitudes where particulates

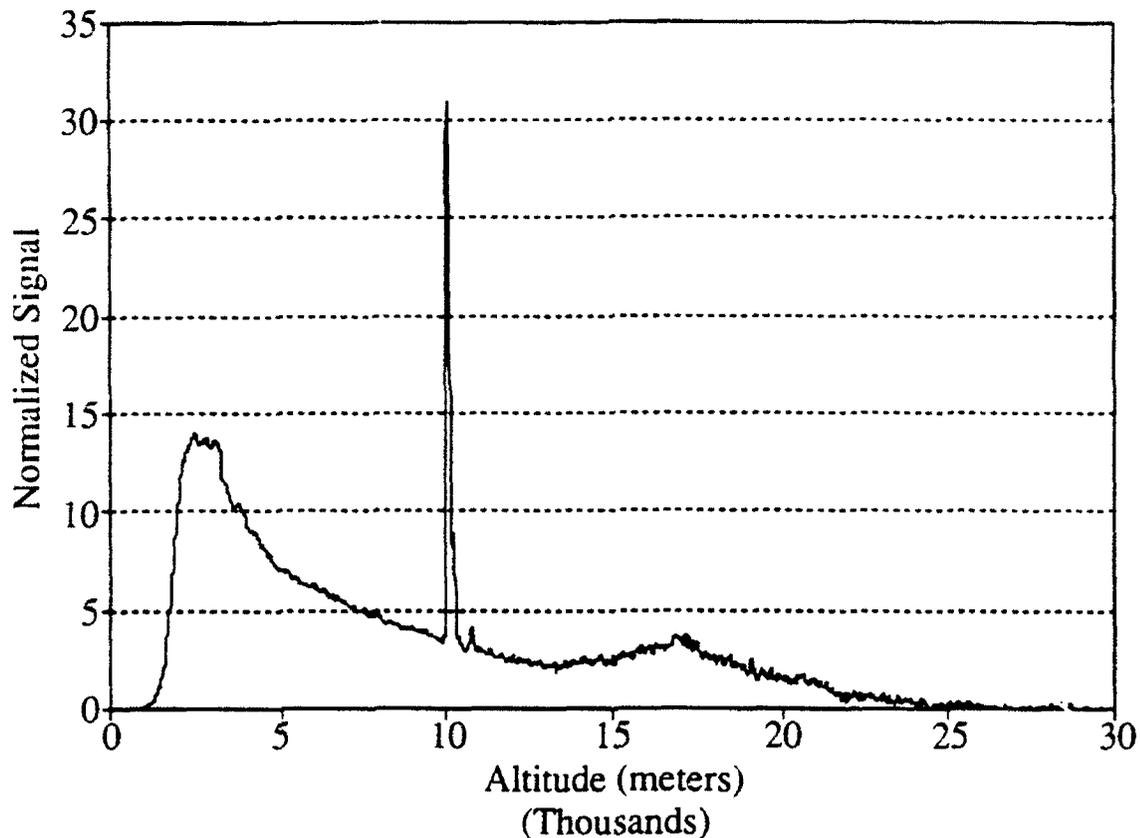


Figure 4. Example of normalized data from August 18, 1992.

exist, the signal is enhanced. Three such altitudes are suggested by the figure: the region just above the beam crossover (boundary layer aerosols); the region at 10 km (cirrus cloud); and the region from 12-24 km (stratospheric aerosol).

The regions of enhanced backscatter can be brought out more clearly by dividing out the backscatter signal from clear air. The result, known as a scattering ratio, is shown in Figure 5. Here the three regions where particulates occur are clearly shown as regions where the scattering ratio is greater than unity. Because the lidar system is not calibrated, it is not possible to know a priori the signal levels which correspond to clear air, so

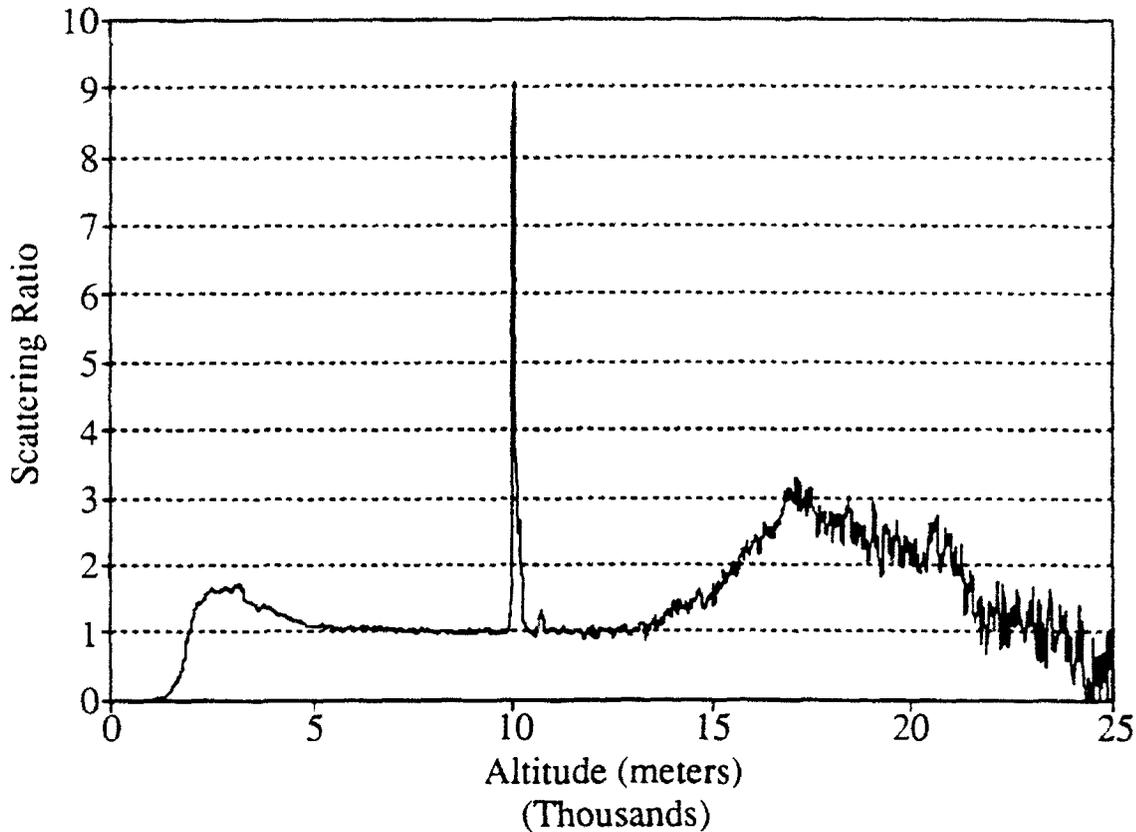


Figure 5. Example of scattering ratio from August 18, 1992.

in practice the scattering ratio is obtained by dividing the normalized data by an air density profile (obtained from the Standard Atmosphere) and normalizing the result to unity in some range of altitude where there appears to be little particulate scattering.

The scattering ratio can also be used to obtain a measure of cloud transmittance, by normalizing the ratio to unity below the cloud and examining it above the cloud. This method is illustrated in Fig. 6. In this example, the value above the cloud is about 0.8. This is a two-way transmittance because the laser light had to propagate up through the cloud to the higher

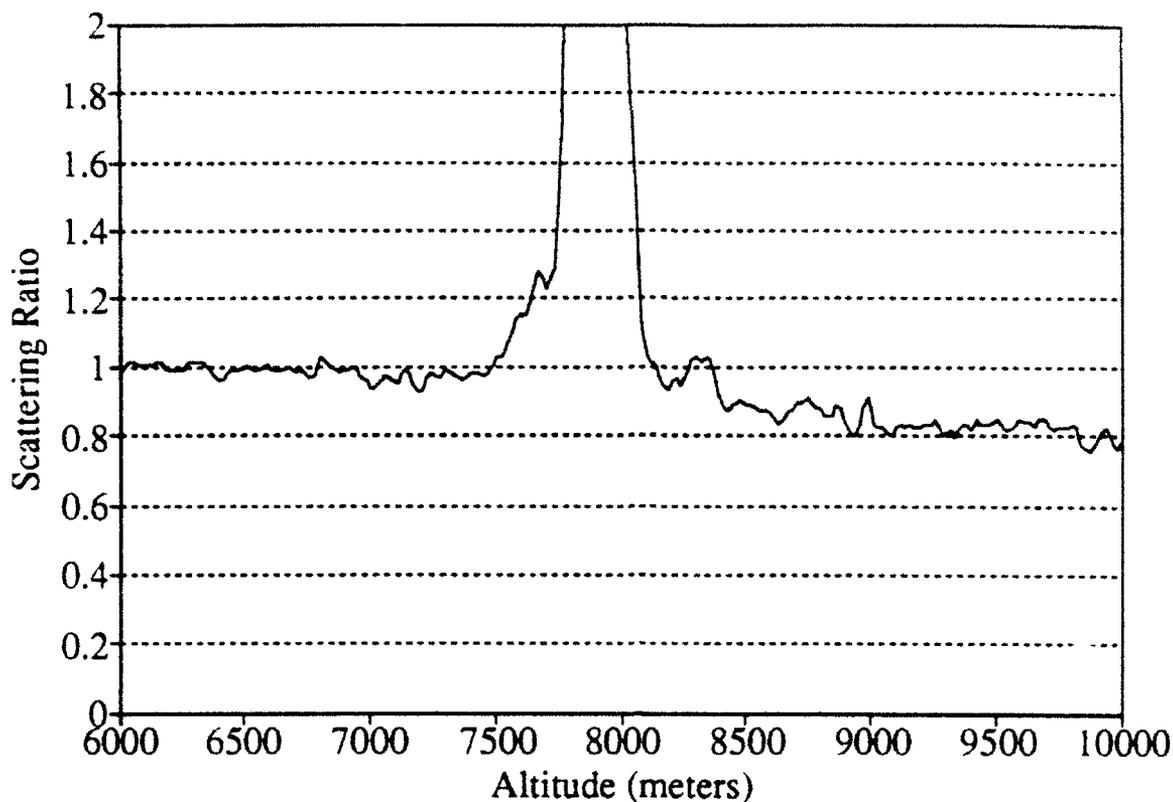


Figure 6. Scattering ratio from August 18, 1992.

altitudes and then back down through the cloud to the lidar receiver. A one-way transmittance of about 0.9 is obtained by taking the square root.

The data shown in the three figures above is one record out of a sequence of 90. These records were obtained by operating the lidar system in a timed acquisition mode, in which data was automatically acquired every 30 seconds for 45 minutes. The complete data set is shown in Fig. 7. This figure shows normalized data in the altitude range 3 to 12 km, with signal increasing vertically. During the time period shown, two cirrus clouds drifted across the zenith. The first cloud had a base

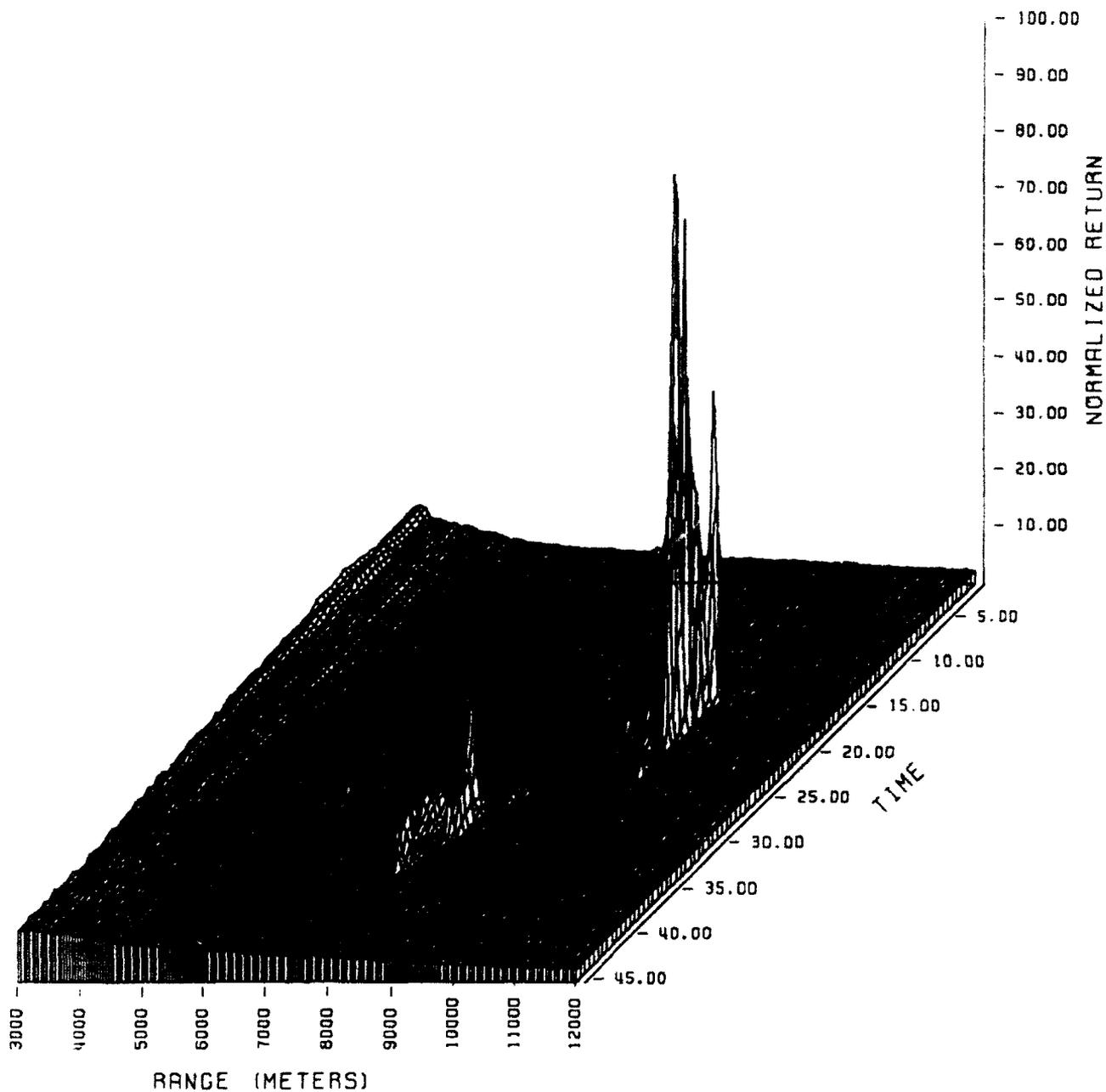


Figure 7. Time sequence of normalized cirrus data.

height of 9.4 km, a thickness of 500 m, and a transmittance of 0.90. The second cloud had a base height of 7.7 km, a thickness of 400 m, and a transmittance of 0.98. These clouds were not visible to rooftop observers, even though the sky was illuminated by bright moonlight.

On August 21, 1992 the 14-inch telescope was mounted on the side of the 1.5 meter telescope and boresighted to it. The lidar receiver was first aligned by using a microwave antenna on a tower two miles away as a convenient boresighting target. Efforts to reliably align the receiver optic axis with the laser beam axis continued until the end of the project.

Alignment usually began during daylight hours, when images of a distant target could be observed in the receiver optical train by projecting them on pieces of white paper. This is especially convenient for focussing. The 14-inch telescope is normally on top of the 1.5 meter, which makes it inaccessible, so the 1.5 meter telescope was flipped over and nearly horizontal for this step in the lidar alignment. The next step was to make small adjustments to alignment while propagating the CVL beam at night and monitoring the lidar signal with an oscilloscope. This was done at 45° elevation, again with the 1.5 meter telescope upside down. Serious alignment errors always occurred when the telescope was flipped over to its normal operating position.

A detailed investigation showed that the alignment problem was caused by flexing of components inside the Celestron telescope under the influence of gravity, and that its optic axis deflects by an angle of about 1 milliradian in going from upside down to rightside up. This deflection is larger than the long-range receiver field of view (see Table 3.) Efforts were made to stiffen the primary mirror mounts and back casting, but the deflection remained at about 1 milliradian. The most successful alignment procedure involved deliberately misaligning the lidar receiver by about 1 mr in the opposite direction so that it would come into proper alignment as the telescope was turned over, but even this technique was not entirely successful because the alignment was not stable from night to night.

We did operate the lidar system simultaneously with stellar photometry on two occasions at the end of the project, but the sky was very clear on these nights and so the data does not demonstrate the usefulness of lidar for monitoring variations in atmospheric transmittance.

CONCLUSIONS

The results shown in Figs. 3-7 show that the prototype lidar system can easily detect subvisual cirrus and that it can be used to monitor changes in sky transmittance to an accuracy of a few percent on a time scale of seconds.

The usefulness of the prototype lidar was seriously limited, however, by alignment instability which was traced to flexing in the receiver telescope. Efforts to reduce this flexing were unsuccessful, so this telescope should be replaced with a more stable one.

ACKNOWLEDGEMENTS

The author would like to thank Dr. Robert Q. Fugate and the entire staff at SOR for their excellent support during the course of the project, particularly Ray Ruane who coordinated and participated in all aspects of system assembly, checkout, alignment, and installation, and Major Leatherman and Captain Morgenstern who participated in all of the lidar measurements and provided invaluable additional support.

DYNAMICS OF THE REACTIONS OF Ar⁺, N₂⁺ AND Kr⁺ WITH NO

Susan T. Graul
Assistant Professor
Department of Chemistry

Carnegie Mellon University
4400 Fifth Avenue
Pittsburgh, Pennsylvania 15213-3890

Final Report for:
Summer Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Hanscom Air Force Base, Massachusetts 01731-5000

September 1992

DYNAMICS OF THE REACTIONS OF Ar⁺, N₂⁺ AND Kr⁺ WITH NO

Susan T. Graul
Assistant Professor
Department of Chemistry
Carnegie Mellon University

Abstract

Absolute cross sections for the reactions of Ar⁺(²P_{3/2}, 1/2), N₂⁺(²Σ), and Kr⁺(²P_{3/2,1/2}) with NO have been measured in a guided-ion beam apparatus for the laboratory energy range of 0.1-20 eV. Throughout this energy regime, charge transfer is the dominant process, and in all three cases, the cross section decreases with increasing collision energy. Preliminary measurements of time-of-flight distributions indicate that the NO⁺ charge transfer products are formed predominantly with near-thermal kinetic energies, but a minor backscattered channel is observed at low energies that can be associated with long-lived complex formation. For all three reactions, the dissociative charge transfer products N⁺ and O⁺ are also formed, with appearance energies close to their thermodynamic thresholds. The Ar⁺ + NO reaction yields ArN⁺ in the energy range between about 5 and 20 eV, but no ArO⁺ could be detected.

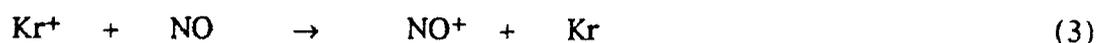
DYNAMICS OF THE REACTIONS OF Ar⁺, N₂⁺ AND Kr⁺ WITH NO

Susan T. Graul

INTRODUCTION

One of the missions of the Phillips Laboratory is to research the interaction of spacecraft effluent gases with the upper atmosphere. In addition to providing information that can aid in design improvements and troubleshooting, such research can assist in interpreting the data from spaceborne experiments by probing the extent to which spacecraft in low-Earth orbit perturb their immediate environment. An example of a local perturbation that arises from contaminants originating with spacecraft is emissions that are observed in the vicinity of the space shuttle. These emissions have been identified as arising from electronically excited NO⁺ ions, which are believed to result from charge transfer reactions between O⁺ (the major ionic constituent at altitudes of 200-400 km [1]) and NO, which outgases from the surfaces of the shuttle. As a result of the orbiting velocity of the shuttle, ion-molecule reactions that take place near the shuttle surfaces can proceed at collision energies ranging from thermal energies to several eV in the center of mass frame of reference. Although the O⁺ + NO charge transfer reaction is strongly exoergic for formation of ground state products, formation of an electronically excited state NO⁺ ion is at least 2 eV endoergic and therefore must involve translational - internal energy transfer at hyperthermal collision energies. Thus, production of excited NO⁺ should be accompanied by a loss of relative kinetic energy of the collision partners, which will be reflected by shifts in the velocity distributions of the NO⁺ products.

In order to gain insight to the dynamics of exoergic charge transfer reactions involving NO, we initiated this summer the study of a series of such reactions. We began our investigations by studying the reactions of Ar⁺, N₂⁺, and Kr⁺ with NO (eqs. 1-3).

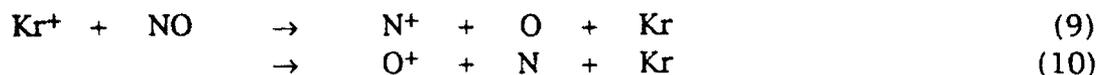
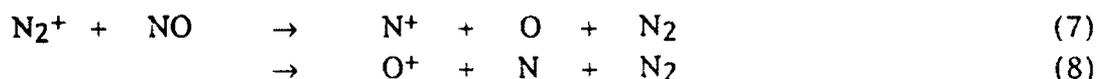
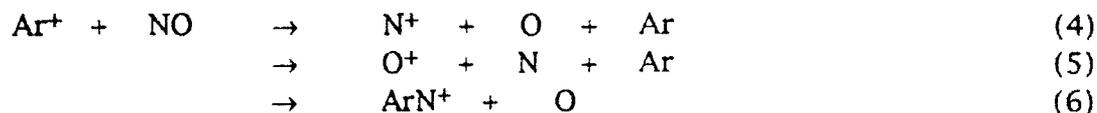


The Ar⁺, N₂⁺, and Kr⁺ reactant ions can be generated in abundance by electron ionization of the corresponding neutral gases, and the resulting high intensity of the reactant ions facilitates the measurement of accurate reaction cross sections and product ion time-of-flight (TOF) distributions.

Like the $O^+ + NO$ charge transfer reaction, each of the reactions in eqs. 1-3 is highly exoergic for formation of ground state products. It has been suggested that the efficiency of charge transfer reactions for atoms and small molecules is controlled by the combined requirements of energy resonance and favorable Franck-Condon factors for the transition from the reactant states to the product states [2-6]. Because reactions 1-3 are strongly exoergic (and hence extremely non-resonant), formation of ground state products would be predicted to be very inefficient. However, at thermal collision energies reactions 1 and 2 proceed with relatively large rate constants that are consistent with reaction occurring on a significant fraction of encounters. Reaction 3 is approximately an order of magnitude slower, but still is much faster than would be predicted based on the extreme non-resonance of the reactant and product ground states and the vanishingly small Franck-Condon factors associated with ionization to high vibrational levels of the ground state NO^+ ion.

It is certain that part of the reason for the apparent failure of the energy resonance and Franck-Condon criteria in determining the probability of charge transfer is the perturbation introduced by the long-range interactions in ion-molecule reactions. What may be more important in the case of the Ar^+/NO and N_2^+/NO reactions is the existence of "near"-resonance with excited states of the products. The reaction of Ar^+ with NO is near-resonant (0.03 eV) for production of $NO^+(a^3\Sigma^+)$. The reaction of N_2^+ with NO is slightly endoergic (ca. 0.15 eV) for formation of $NO^+(a^3\Sigma^+)$, and exoergic by only about 0.15 eV for $NO^+(X)$ and $N_2(a^3\Sigma^+)$. These asymptotic energy gaps are comparable to ion-induced dipole interaction energies, which at close range can act to narrow the gaps further. The average product ion kinetic energies measured for reaction 1 and 2 at near-thermal collision energies in an ion cyclotron resonance spectrometer indicate that very little energy is partitioned to relative translation of the products, suggesting that the products are internally excited [7]. Evidence from beam experiments suggests that $NO^+(a)$ (rather than $NO^+(X, v>0)$) is the major product for reaction 1 at thermal energies [8]. The nonzero (but nonetheless low) efficiency of the $Kr^+ + NO$ charge transfer [9] may be a consequence of modification by the long-range potential of the unfavorable Franck-Condon factors for production of the highly vibrationally excited $NO^+(X)$ ion that would be the only product state near-resonant with the Kr^+/NO reactants.

We have examined the energy dependence of the cross sections for reactions 1-3 over the energy range of 0.1 to about 20 eV in the center-of-mass frame of reference. In addition, cross sections were measured for the reaction channels that lead to O⁺ and N⁺ for each reaction, and to formation of ArN⁺ from the reaction of Ar⁺ with NO (eqs. 4-10).



For reactions 1 and 2, NO⁺ product ion TOF distributions were collected at several interaction energies. In this report, selected preliminary data from these investigations is presented. Although the studies are not complete, it is possible to draw some conclusions about the mechanisms of the reactions and the internal energies of the products formed.

EXPERIMENTAL

The cross section and time-of-flight measurements were carried out using a guided-ion-beam apparatus, in which ions in the interaction region are confined by a radiofrequency field. The base apparatus is a tandem mass spectrometer, which has been described previously [10], in which the collision cell and associated ion optics have been replaced by an octopole beam guide [11,12] (Fig. 1). The instrument will be described briefly here.

The Ar⁺, N₂⁺, and Kr⁺ ions are produced by 25 eV electron ionization of argon, nitrogen or krypton. The ions are accelerated and mass-selected in a Wien velocity filter. Following mass selection, the ions are decelerated and injected into a 7.36 cm long octopole ion guide of 6 mm inner radial diameter and 2 mm rod diameter. A 10 MHz radiofrequency potential with typical amplitude of 90 V r.m.s. is applied to the octopole rods, and acts to trap the ions in the radial direction without significantly altering the axial velocity distribution. The octopole passes through a 3.5 cm long

GUIDED-ION BEAM EXPERIMENT

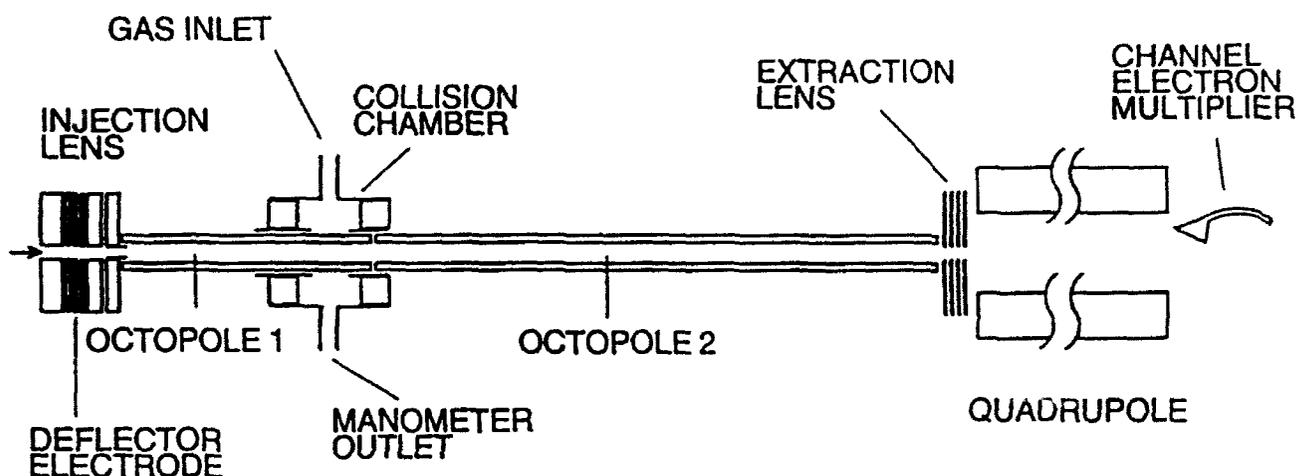


Figure 1. Schematic of octopole beam guide.

collision cell containing neutral NO at pressures of 150-250 μ torr (measured with a Baratron capacitance manometer). At the end of the collision cell, the first octopole connects to a second octopole of equal radial dimensions, but 16.7 cm in length. The second octopole is normally held at a potential some 350 mV lower than the first to aid in extracting slow ions formed in the collision cell. Potentials applied to cylindrical electrodes at the entrance and exit of the collision cell further aid in collecting ions that are back-scattered in the lab frame. At the terminus of the second octopole, the ions are extracted and focused into a quadrupole mass analyzer, and detected with a channelplate electron multiplier.

The primary ion energy and the width of the energy spread are determined by performing retardation potential scans of the octopole bias for a constant injection potential. Minimum energy half-widths are typically obtained by tuning the injection optics at the entrance of the octopole such that the ions are injected with less than 1 eV axial energy. Typical half-widths observed in these experiments were 250-350 meV. Average ion energies can also be determined by pulsing the primary

ion beam and analyzing the flight time difference measured with different bias potentials on the second octopole. The energies so measured typically differed by less than a few hundred meV from the nominal ion energy for the same octopole potential as measured by retardation scans, with the small differences arising from effects of the gate pulse on the ion energy.

Integral cross sections are calculated from the relative integrated intensities of the product ions and the reactant ion as measured under single-collision conditions. For measurements carried out at low conversion, the intensities are used in eq. 11:

$$\sigma = \frac{I_p}{[\Sigma I_p + I_R] n l} \quad (11)$$

in which I_p and I_R are the integrated intensities of the product and reactant ions respectively, n is the number density of the neutral reactant, and l is the reaction path length. At each collision energy, the extraction optics are tuned for optimal ion collection of both reactant and product ions. The absolute error of the present cross section measurements is estimated at approximately 30%, whereas the relative error is probably considerably smaller.

For the measurement of product or reactant ion time-of-flight (TOF) distributions, the primary ion beam is deflected out of the collision cell by applying an approximately 10 V potential to the deflector electrode (Fig. 1). A triangular 15-20 V pulse is used to drop this gate and allow a burst of reactant ions into the collision cell. The effective width of this pulse is usually less than 5 μ sec, whereas the product TOF distributions were typically more than 250 μ sec wide. The gate pulse simultaneously sends a start signal to the detector timer (which receives a stop signal when an ion is detected at the channelplate multiplier) and triggers the function generator that supplies the 10 MHz radiofrequency potential to the octopole guide. For a repetition rate of 3 KHz, the rf field was kept on for about 300 μ sec, then shut off to allow any remaining ions in the octopole guide to escape before the next timing cycle.

Significant background ion signals corresponding to NO^+ were observed when the reactant ions were deflected completely out of the collision cell. The background NO^+ signal was highest for the Ar^+ and Kr^+ reactants and was negligible for N_2^+ . These NO^+ ions probably arise from Penning ionization of NO by fast metastable neutral Ar or Kr atoms, which most likely result from neutralization of the accelerated and focused Ar^+ or Kr^+ ions as they skim metal surfaces in the Wien filter

or injection optics. Corrections to the NO^+ product TOF distributions were made by subtracting the background TOF distribution collected with the ions always deflected out of the cell from the distributions collected with the ions pulsed into the cell. For cross section measurements, in which the reactant ion beam is continuous rather than pulsed, the contribution to the product NO^+ signal by Penning ionization of NO is less than a few percent and can be neglected.

RESULTS

Known energetics for the reactions of Ar^+ , N_2^+ , and Kr^+ with NO are listed in Table 1.

$\text{Ar}^+ + \text{NO}$: The energy dependence measured for the $\text{Ar}^+ + \text{NO}$ charge transfer reaction for 0.1 - 20 eV collisions (center of mass) is shown in Fig. 2. Below about 6 eV, the cross section follows an approximately $E^{-1/2}$ dependence, but at higher energies the cross section decreases more rapidly with increasing energy. An $E^{-1/2}$ dependence is the predicted energy dependence for the cross section of a ion-molecule reaction that involves a long-lived intermediate complex.

At collision energies above about 5 eV, three products besides NO^+ are observed, corresponding to dissociative charge transfer and atomic ion transfer (eqs. 4-6). The appearance energies for these products is within 0.5 eV of the thermochemical thresholds (Table 1). The energetic onsets for these products correspond approximately to the break in the cross section for charge transfer (Fig. 3).

Time-of-flight distributions were measured for the NO^+ products from charge transfer at several collision energies. In the interest of brevity, these distributions are not shown here, but the results will be summarized briefly. At all collision energies, the major portion of the NO^+ product ions are formed with near-thermal velocities. At the lowest collision energies (2-5 eV), a significant fraction of fast NO^+ ions are also formed, indicative of back-scattering in the lab frame. The fraction of back-scattered NO^+ ions decreases as collision energy increases.

Table 1. Energetics for the reactions of Ar⁺, N₂⁺, and Kr⁺ with NO.

Reactants	Products	ΔE , eV
Ar ⁺ + NO	Ar + NO(X ¹ Σ^+)	-6.50
	Ar + NO(a ³ Σ^+)	-0.03
	Ar + NO(A ¹ Π)	+2.61
	Ar + N + O ⁺	+4.40
	Ar + N ⁺ + O	+5.32
	Ar ⁺ + N + O	+6.50
	ArO ⁺ + N	+3.66
	ArN ⁺ + O	+4.98
N ₂ ⁺ + NO	N ₂ + NO(X)	-6.32
	N ₂ + NO(a)	+0.15
	N ₂ + NO(A)	+2.79
	N ₂ + N + O ⁺	+4.58
	N ₂ + N ⁺ + O	+5.50
	N ₂ ⁺ + N + O	+6.50
	N ₂ (A ³ Σ_u^+) + NO ⁺ (X)	-0.15
	N ₂ (A) + NO ⁺ (a)	+6.62
Kr ⁺ + NO	Kr + NO ⁺ (X)	-4.74
	Kr + NO ⁺ (a)	+1.73
	Kr + NO ⁺ (A)	+4.37
	Kr + N + O ⁺	+6.16
	Kr + N ⁺ + O	+7.08
	Kr ⁺ + N + O	+6.50

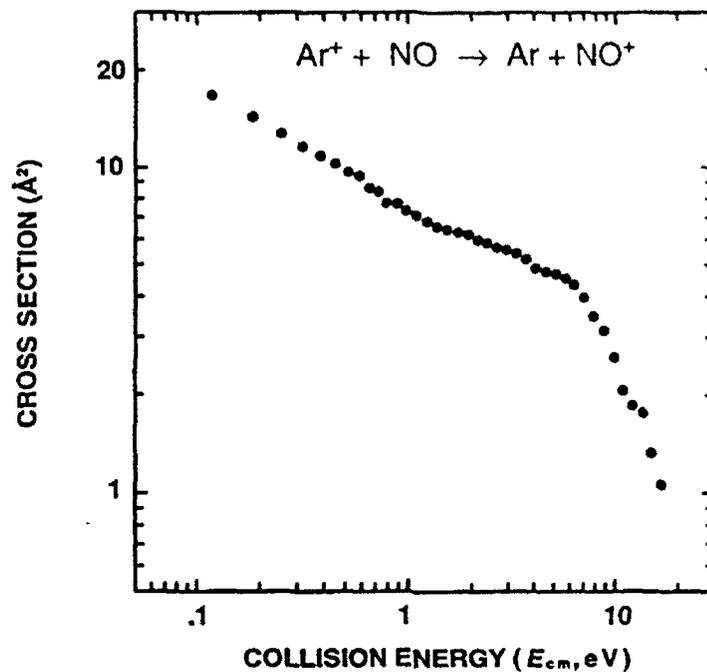


Figure 2. Absolute cross sections as a function of collision energy for the charge transfer reaction of Ar^+ and NO .

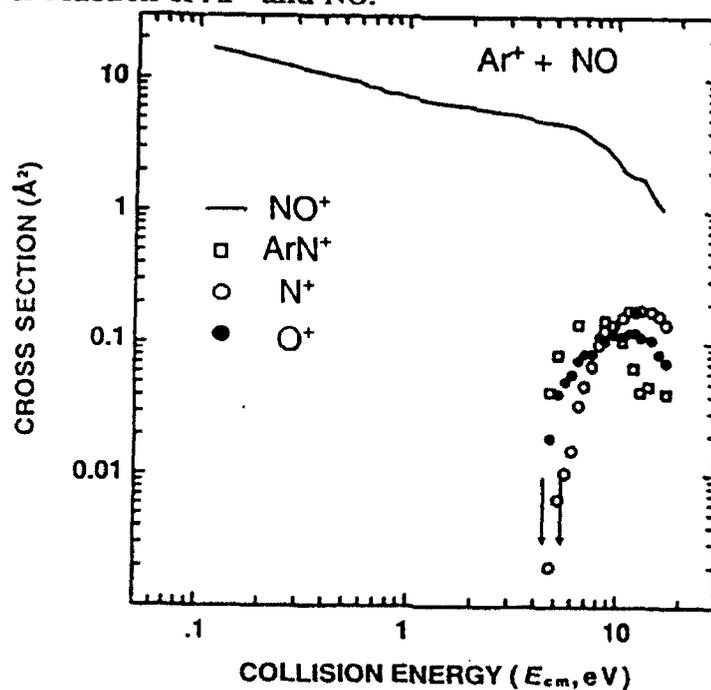


Figure 3. Absolute cross sections for production of NO^+ , ArN^+ , N^+ , and O^+ from the reaction of Ar^+ and NO . Vertical arrows mark the thermochemical threshold energies for O^+ and N^+ .

$N_2^+ + NO$: The charge transfer reaction of N_2^+ with NO to form ground state NO^+ is only slightly less exoergic than the Ar^+/NO reaction (Table 1). Two "near"-resonant charge transfer channels exist: the one leading to $NO^+(a^3\Sigma_u^+)$ and ground state N_2 is about 0.15 eV endoergic, and a second leading to $N_2(a^3\Sigma_u^+)$ and $NO^+(X)$ is about 0.15 eV exoergic. In spite of the larger energy gaps for the N_2^+/NO reaction compared to the Ar^+/NO reaction, the $N_2^+ + NO$ charge transfer cross sections are consistently larger (Fig. 4). The opposite would be expected based on energy resonance criteria. There also appears to be some structure in the N_2^+/NO cross section, in particular near the break point at about 10 eV.

The fragment ions N^+ and O^+ are observed for the N_2^+/NO reaction at collision energies above 5 eV (Fig. 5). The appearance energies for these products agree within 0.5 eV with the thermochemical thresholds (Table 1). In the absence of isotope labeling, collision-induced dissociation of N_2^+ (eq. 12) cannot be ruled out as a source for the N^+ ions.



However, the energy required for CID of N_2^+ is nearly 3 eV greater than that required for formation of N^+ by dissociative charge transfer, and the cross section for N^+ is a smooth function of collision energy even above the CID threshold. This suggests that the major (if not exclusive) source of N^+ is dissociative charge transfer to NO. Although ion signals were observed that may correspond to N_3^+ and N_2O^+ , they were extremely weak, with cross sections less than 0.01 \AA^2 .

The time-of-flight distributions for the NO^+ product of charge transfer are qualitatively similar to those observed for the $Ar^+ + NO$ reaction, in that the major fraction of the NO^+ product ions are formed with near-thermal velocities.

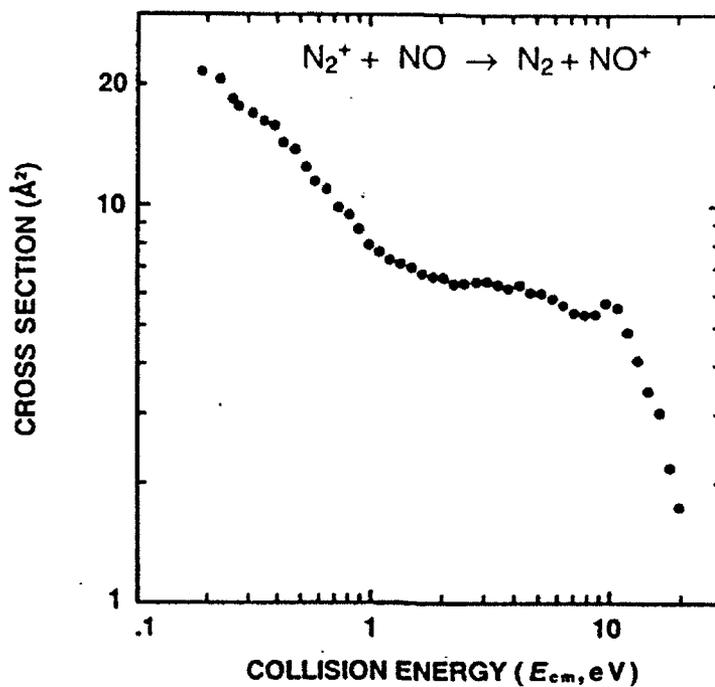


Figure 4. Absolute cross sections as a function of collision energy for the charge transfer reaction of N_2^+ and NO.

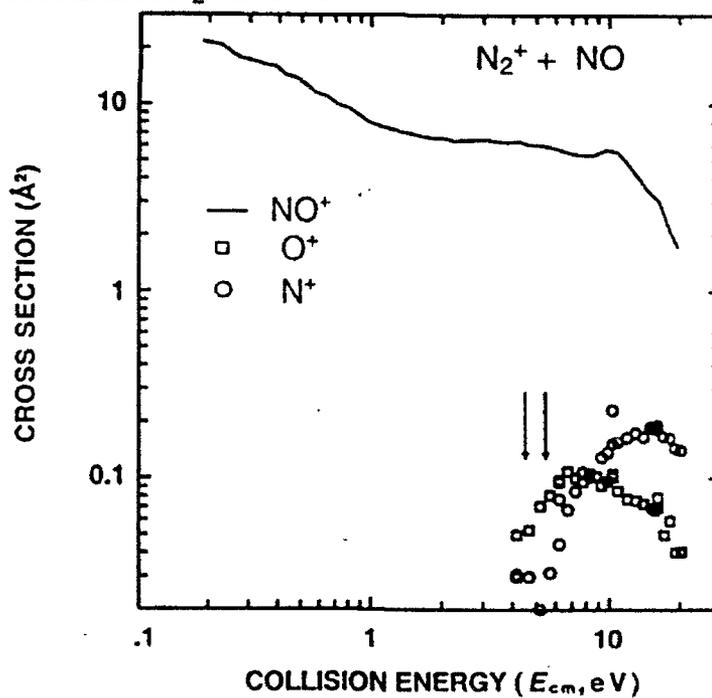


Figure 5. Absolute cross sections for production of NO^+ , N^+ , and O^+ from the reaction of N_2^+ and NO. The vertical arrows indicate the thermochemical threshold energies for O^+ and N^+ .

$\text{Kr}^+ + \text{NO}$: This charge transfer reaction is strongly exoergic to form $\text{NO}^+(\text{X})$ (Table 1), but unlike the Ar^+/NO and N_2^+/NO reactions, there are no product electronic states in near-resonance with the reactants. The charge transfer reaction has a much smaller cross section than either Ar^+/NO or N_2^+/NO , although the energy dependence (shown in Fig. 6) is similar to that observed for Ar^+/NO . The fragment ions N^+ and O^+ appear at collision energies equal to and greater than their thermodynamic thresholds. Formation of O^+ is favored over N^+ at all energies above 4 eV, and the cross sections for these products are nearly an order of magnitude smaller than for the Ar^+/NO and N_2^+/NO reactions.

The small cross sections for charge transfer in the Kr^+/NO reaction precluded measurement of product ion TOFs within the time available under the Summer Research Program.

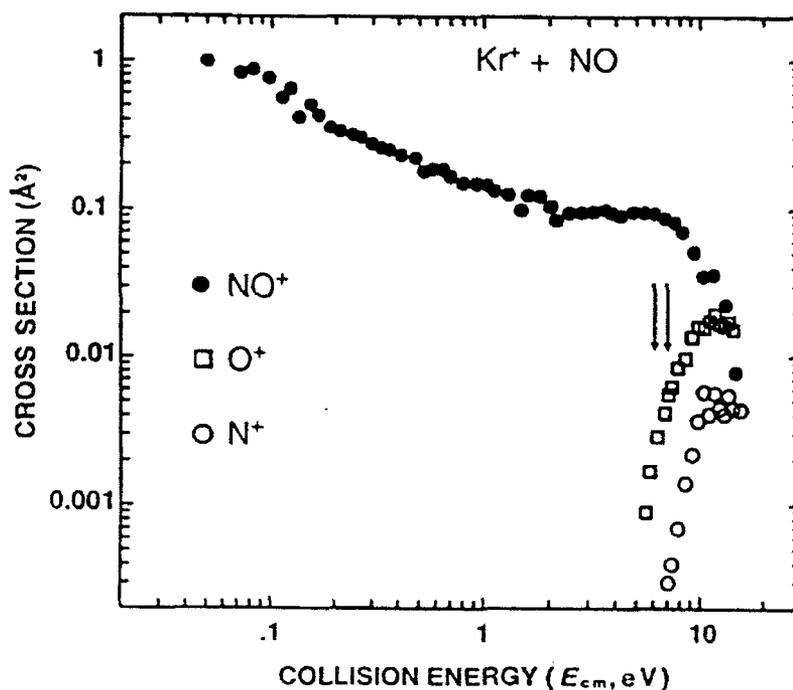


Figure 6. Absolute cross sections for production of NO^+ , N^+ , and O^+ from the reaction of Kr^+ and NO . The vertical arrows indicate the thermochemical threshold energies for O^+ and N^+ .

DISCUSSION

Kobayashi measured the cross sections for the charge transfer reactions of Ar^+ and Kr^+ with NO in a drift tube in the collision energy range of 0.04 - 3 eV [9]. For the $\text{Ar}^+ + \text{NO}$ reaction, he reported charge transfer cross sections that are 25-50% larger than the present results, and observed a shallow minimum or plateau in the charge transfer cross section above 1 eV collision energy (center of mass). The differences between Kobayashi's results for $\text{Ar}^+ + \text{NO}$ and the present measurements are within the combined experimental errors. However, in the case of $\text{Kr}^+ + \text{NO}$, the differences are considerably larger. At the lowest energies (ca. 0.1 eV), our cross sections are about twice those measured by Kobayashi, and we do not observe the prominent minimum he reported at 0.7 eV collision energy. The reason for these discrepancies is not clear. Ng and coworkers have demonstrated that individual spin-orbit states of the ground state Ar^+ ion can have markedly different reactivity [13], but there is no reason to expect that the Ar^+ and Kr^+ ions formed in our experiments have significantly different populations of spin-orbit states from those formed in Kobayashi's experiments; in both systems, the ions are generated at low pressure by electron ionization (albeit at different electron energies). It is possible, however, that the upper spin-orbit state is quenched in the drift region of Kobayashi's apparatus, in which case he observes primarily the reactivity of the lower spin-orbit state. Another possible interference in these experiments is Penning ionization of NO by metastable electronically excited state Ar or Kr atoms. We checked for ionization by metastable Ar or Kr atoms by deflecting the ion beam out of the octopole and monitoring the background NO^+ signal. For cross section measurements, the background NO was only a few percent of the charge transfer reaction product signal. Therefore, contributions to the NO^+ signal from Penning ionization should be negligible in our experiments.

The time-of-flight distributions for the NO^+ products formed by charge transfer from Ar^+ or N_2^+ are not shown in the report, but a qualitative summary of the results has been included. For both reactions, most of the NO^+ products are formed with near-thermal velocities, suggestive of charge transfer at long range with very little momentum transfer. However, some fraction of the NO^+ ions have short flight times, corresponding to large velocities. These ions could result from low-impact-parameter collisions or from a long-lived collision complex in which momentum transfer can occur. The fraction of the "fast" NO^+ ions decreases as

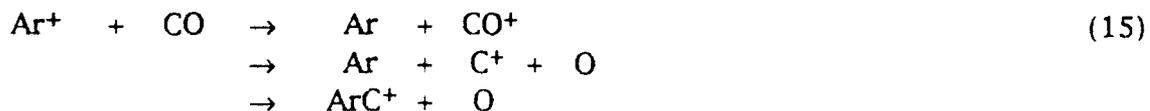
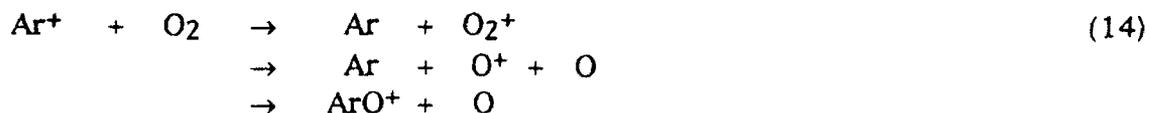
collision energy increases, just as the lifetime of a collision complex (and consequently the extent of momentum transfer) can be expected to decrease with increasing collision energy. It is clear from the experimental velocity distributions for charge transfer with N_2^+ and Ar^+ that most of the charge transfer reactions occur with very little internal \leftrightarrow translational energy transfer. This would support the formation of $NO^+(a^3\Sigma^+)$, or, for the N_2^+/NO reaction, $N_2(a^3\Sigma_u^+) + NO^+(X)$, in the near-resonant charge transfer channels as the major process throughout the collision energy range studied.

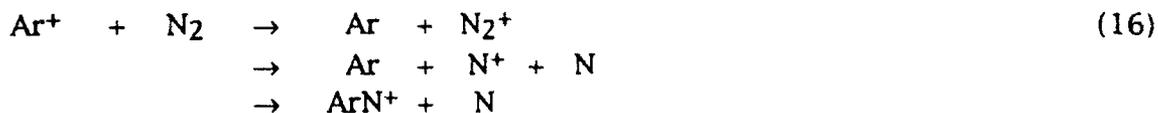
Reaction channels that do require translation \rightarrow NO internal energy transfer are the formation of N^+ and O^+ , which are relatively minor processes in these reactions. In the Ar^+/NO and Kr^+/NO reactions, the collision energy at which a break occurs in the charge transfer cross section corresponds approximately to the appearance energy for N^+ and O^+ , suggesting that the break might be correlated with the two new products. However, in the N_2^+/NO reaction, the break comes nearer to the energy at which the N^+ and O^+ cross sections maximize. In none of the reactions are the N^+ and O^+ cross sections large enough to compensate for the sudden drop in the charge transfer cross section at collision energies above 10 eV. An explanation that would be consistent with the formation of only (or predominantly) $NO^+(a)$ throughout the energy range, as suggested by the velocity distributions, is that the collisions are resulting in formation of N and O atoms via collisionally activated dissociation of NO (eq. 13).



The thermochemical threshold for this process is 6.497 eV, and the breaks in the charge transfer cross sections occur at energies about equal to or greater than this lower limit.

The onsets of formation of N^+ and O^+ products occur near the thermochemical thresholds for dissociative charge transfer. Analogous products were observed by Ng and coworkers in the reactions of Ar^+ with O_2 [13], CO [14,15], and N_2 [16] (eqs. 14-16).





Based on their analysis of the product ion retardation potential curves, they concluded that the dissociative charge transfer products arise via predissociation of nascent electronically excited charge transfer products (O_2^+ , CO^+ , or N_2^+). These reactions also yielded ArO^+ , ArC^+ , and ArN^+ , which Ng suggested arise by a charge transfer predissociation mechanism. This is apparently a selective process for $\text{Ar}^+ + \text{CO}$, which yields ArC^+ , but not ArO^+ . It seems plausible that similar interpretations will be valid for the reactions of Ar^+ , Kr^+ , and N_2^+ with NO , but further study is required to determine if this is correct.

It is interesting that the Ar^+/NO , Kr^+/NO , and N_2^+/NO systems all show similar reactivity. Although the magnitudes of the cross sections vary considerably, the energy dependences are quite similar, and the higher energy N^+ and O^+ product ions are observed in each case at the thermochemical thresholds. These results suggest that the dynamics of these reactions may be relatively independent of energy resonance criteria.

CONCLUSIONS

The reactions of Ar^+ , N_2^+ , and Kr^+ with NO proceed predominantly via charge transfer at fairly long range and with little momentum transfer. The energy dependences of the cross sections are similar for each of these reactions, but the magnitude of the cross sections vary. Minor reaction channels lead to N^+ and O^+ products from dissociative charge transfer, occurring near the thermochemical threshold.

ACKNOWLEDGMENTS

Coworkers on this project were Skip Williams (Stanford University) and Rainer Dressler (Phillips Laboratory). I am very grateful to Skip and Rainer for an enjoyable and instructive tour.

REFERENCES

1. Jursa, A., Ed. "Handbook of Geophysics and the Space Environment", ADA 167000, 1985.
2. Rapp, D.; Francis, W. E. *J. Chem. Phys.* 1962, 37, 2631.
3. Laudenslager, J. B.; Huntress, Jr., W. T.; Bowers, M. T. *J. Chem. Phys.* 1974, 61, 4600.
4. Campbell, F. M.; Browning, R.; Latimer, C. J. *J. Phys. B* 1980, 13, 4257.
5. Kato, T.; Tanaka, D.; Koyano, I. *J. Chem. Phys.* 1982, 77, 834.
6. Bauer, E.; Fisher, E. R.; Gilmore, F. R. *J. Chem. Phys.* 1969, 51, 4173.
7. Parent, D.; Derai, R.; Mauclaire, G.; Heninger, M.; Marx, R.; Rincon, M. E.; O'Keefe, A.; Bowers, M. T. *Chem. Phys. Lett.* 1985, 117, 217.
8. Herman, Z.; Pacák, V.; Yench, A. J.; Futrell, J. *Chem. Phys. Lett.* 1976, 37, 329.
9. Kobayashi, N. *J. Phys. Soc., Japan* 1974, 36, 259.
10. Dressler, R. A.; Gardner, J. A.; Salter, R. H.; Wodarczyk, F. J.; Murad, E. *J. Chem. Phys.* 1990, 92, 1117.
11. Gerlich, D. in *State-Selected and State-to-State Ion Molecule Reaction Dynamics*, *Adv. Chem. Phys.* vol. 82, part 1; Baer, M., Ng, C. Y., Eds.; Wiley, New York, 1992.
12. Teloy, E.; Gerlich, D. *Chem. Phys.* 1974, 4, 417.
13. Flesch, G. D.; Nourbakhsh, S.; Ng, C. Y. *J. Chem. Phys.* 1990, 92, 3590.
14. Flesch, G. D.; Nourbakhsh, S.; Ng, C. Y. *J. Chem. Phys.* 1991, 95, 3581.
15. Flesch, G. D.; Ng, C. Y. *J. Chem. Phys.* 1988, 89, 3381.
16. Flesch, G. D.; Ng, C. Y. *J. Chem. Phys.* 1990, 92, 2876.

**RADAR SIMULATION AND ANALYSIS TOOLS
FOR SATELLITE ANALYSIS - A RECOMMENDATION**

**Dr. J.M. Henson
Assistant Professor
Department of Electrical Engineering**

**University of Nevada
Reno, NV**

**Final Report for:
Summer Research Program
Phillips Laboratory**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.**

August 1992

**RADAR SIMULATION AND ANALYSIS TOOLS
FOR SATELLITE ANALYSIS - A RECOMMENDATION**

**Dr. J.M. Henson
Assistant Professor
Department of Electrical Engineering
University of Nevada**

Abstract

In order for the Phillips Laboratory, Advanced Weapons Directorate, Satellite Signature and Imaging (WSAI) working group to realize its mission, certain radar data and image simulation and analysis tools and techniques must be defined and integrated into the group's existing modeling and analysis techniques. While there are several modeling issues (multispectral model compatibility requirements) that are of serious concern to WSAI, these concerns are beyond the scope of this report. Rather, it is our purpose herein to briefly describe current radar related data analysis, imaging, image analysis and software simulation options and to recommend that WSAI adopt the Lincoln Laboratories RCS Toolbox and the Enhanced Lincoln Interactive Analysis System (ELIAS) software packages to support its signature and imaging simulation and analysis requirements.

RADAR SIMULATION AND ANALYSIS TOOLS FOR SATELLITE ANALYSIS - A RECOMMENDATION

J.M. Henson

INTRODUCTION

The basic radar theory, associated hardware, and signal processing analysis techniques that underlie the current technologies appear, based on the literature, to have been fully developed and operational by the mid 1970's. It is also clear, based on the literature, that Lincoln Labs has been and continues to be the primary developer of the technology. Indeed, at least two of our primary radar sites, ARPA - Lincoln C-Band Observables Radar (ALCOR) and HAYSTACK, are designed by Lincoln Laboratories (LL). It is not surprising then that the same laboratory has also developed primary radar data analysis procedures and algorithms that currently define satellite analysis -- with respect to the radar return. The LL **RCS Toolbox** and **Enhanced Lincoln Interactive Analysis System (ELIAS)** software packages are the inevitable results of such an extensive involvement in the technology.

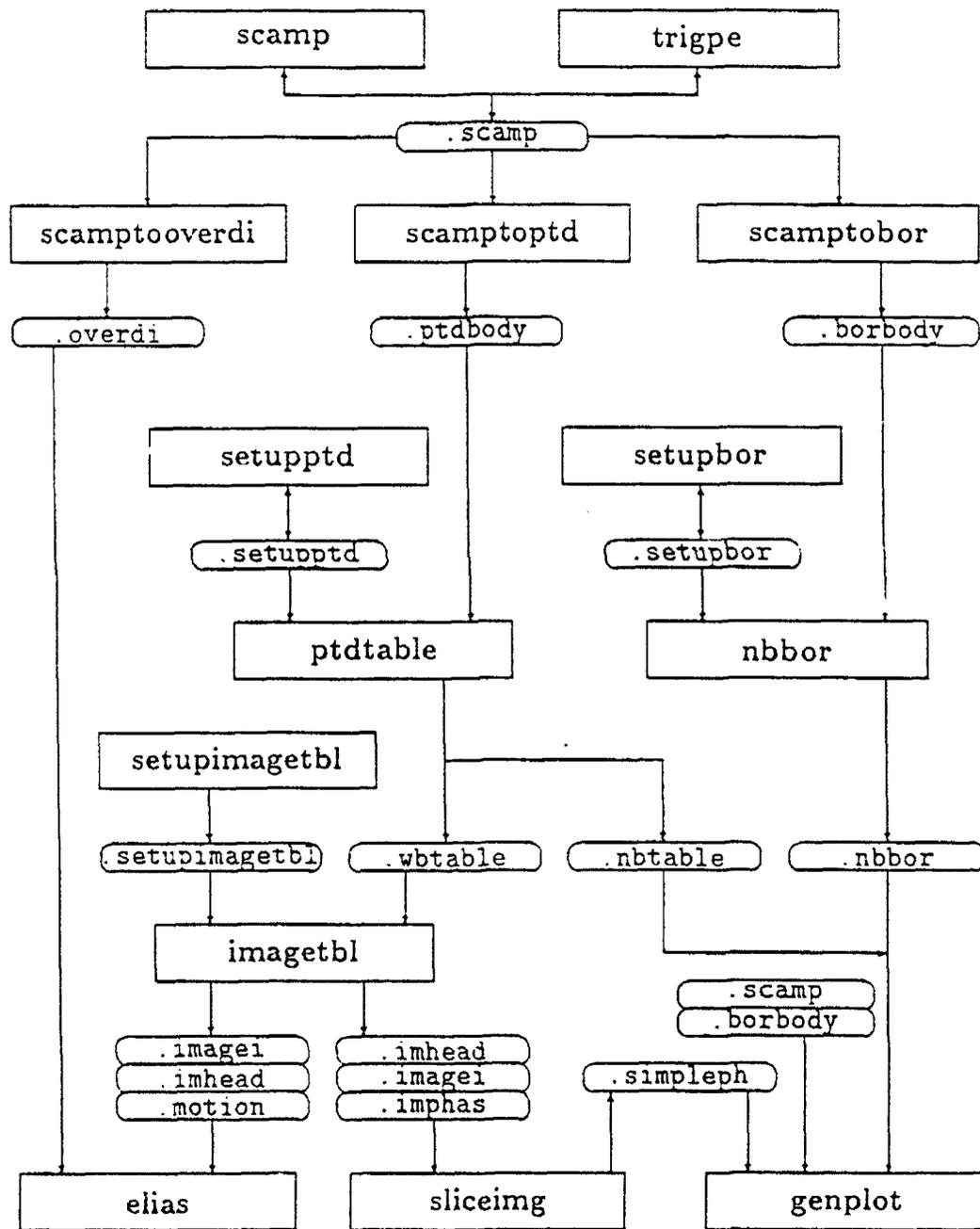
Other software options (including **SRIM**, developed by the Environmental Research Institute of Michigan) simply are not intended or geared to simulate real world satellite imaging and analysis techniques. In fact, it appears that no other software simulations are available that even begin to offer the capabilities of **RCS Toolbox** and **ELIAS**.

Given that there are no suitable alternatives to the LL software products, WASI is faced with acquiring the technical expertise to both use and evaluate the LL products with respect to WASI goals and requirements.

The following section is intended to provide WASI with a brief discussion and evaluation of the LL software based on our own brief experience with it, and further to suggest a general plan by which WASI may develop the radar related expertise and simulation capabilities it requires.

FURTHER RECOMMENDATIONS

Figure 1 (taken from the preliminary version of the LL **User's Guide to RCS Toolbox, version 2.0**) is a block diagram of the 14 programs that comprise the "toolbox". While this is a set of stand alone



programs, there are nevertheless cross dependencies (and some redundancies) with respect to required user inputs. Unfortunately, the current version of the documentation fails to explain these dependencies. Further, the documentation does not describe fully or adequately the input variables that the user must select and for which the user must assign values. Additionally, the documentation does not correlate the nature of these inputs with radar parameters or the nature of software outputs.

The toolbox programs in conjunction with ELIAS, which provides both narrow and wideband data and imagery display and analysis tools does, however, appear to be complete in the sense that full satellite motion simulations as well as narrowband and wideband signature and image analyses can be carried out without the requirement of real data. Unfortunately, nowhere in the documentation can a new user, unfamiliar with current analysis techniques, find instructive information or examples. Given that these software tools have a long developmental history and have been funded and supported by DOD labs other than Phillips Laboratory, we should not and perhaps cannot register a very loud complaint.

Rather, what is necessary at this point is a concerted and sustained effort to understand and work effectively with these products. We would suggest then, that WSAI begin by placing qualified personnel alongside actual analysts who use these software tools to solve real problems on a regular basis. Personnel at FASTC appear willing to share their knowledge and expertise. In this way WSAI may in a highly cost effective manner begin its own education and training in the area of radar satellite imaging, simulation and analysis.

OPTICAL AND ATMOSPHERIC TURBULENCE

submitted by

MAYER HUMI, Professor
DEPARTMENT OF MATHEMATICAL SCIENCES
Worcester Polytechnic Institute
Worcester, MA 01609-2280

with

MARY AGNES BEALS
DEPARTMENT OF MATHEMATICS
Virginia Polytechnic Institute
Blacksburg, VA 24061-0123

Final Report for
Summer Research Program
Phillips Laboratory

Sponsored by
AFOSR, Bolling Air Force Base
Washington, D.C.

September 1992.

OPTICAL AND ATMOSPHERIC TURBULENCE

MAYER HUMI, Professor
DEPARTMENT OF MATHEMATICAL SCIENCES
Worcester Polytechnic Institute

Abstract

For many important applications it is imperative to estimate correctly the spectral density of atmospheric turbulence. To this end one must collect meteorological data and "detrend" it to obtain the turbulent residual. The objective of this project was to experiment with different detrending (= filtering) strategies and gauge their impact on the computed spectral densities of the flow variables. To help accomplish this objective a general purpose software package was written and used.

I. Introduction.

It is a well known fact that the twinkling of the stars is caused by the fluctuations in the index of refraction of the earth atmosphere. These are due to atmospheric turbulence which in turn cause random fluctuations in the effective optical path of the electromagnetic signal. Due to these fluctuations there will be a spreading of the light beam, scintillations and a decrease in the spatial and temporal coherence. The same phenomena affect laser and radar systems and might degrade severely their performance. To compensate for these effects a careful study must be made on the relationship between the meteorological state of the atmosphere and the fluctuations in the index of refraction.

II. The structure constant C_n^2 .

To study the effects of the fluctuations of the index of refraction in a clear atmosphere let $n(\mathbf{r},t)$ denote the index of refraction. This can be written as

$$n(\mathbf{r},t) = \langle n(\mathbf{r},t) \rangle + n_1(\mathbf{r},t) \quad (2.1)$$

where $\langle n(\mathbf{r},t) \rangle$ represents the average part and n_1 the fluctuation. For clear dry air $\langle n \rangle \cong 1$. The fluctuating part n_1 consists of real and imaginary components

$$n_1 = n_{1r} + in_{1i} \quad (2.2)$$

From a practical point of view [1] the most important quantity affecting the signal propagation is the spatial correlation function for n_{1r} , i.e.

$$Bn(\mathbf{r},\boldsymbol{\rho}) = \langle n_{1r}(\mathbf{r})n_{1r}(\mathbf{r} + \boldsymbol{\rho}) \rangle \quad (2.3)$$

Here we neglected the dependence of n_{1r} on t since the time of propagation of an

electromagnetic signal in the atmosphere is short. If we assume further that the atmosphere is spatially homogeneous then $B_n(\mathbf{r}, \rho)$ is independent of \mathbf{r} . Using first order differentials to approximate n_{1r} we have (remember $\langle n_{1r} \rangle = 0$)

$$n_{1r} \cong \frac{\partial n_{1r}}{\partial T} dT + \frac{\partial n_{1r}}{\partial Q} dQ + \frac{\partial n_{1r}}{\partial P} dP \quad (2.4)$$

where T, Q, P stand for the temperature, humidity and pressure. However usually the pressure term is negligible and therefore we can write

$$n_{1r} \cong A_T \frac{dT}{\langle T \rangle} + A_Q \frac{dQ}{\langle Q \rangle} \quad (2.5)$$

Substituting this in eq. (2.3) leads to

$$B_n(\mathbf{p}) = \frac{A_T^2}{\langle T \rangle^2} B_T(\mathbf{p}) + \frac{A_Q^2}{\langle Q \rangle^2} B_Q(\mathbf{p}) + \frac{2A_T A_Q}{\langle T \rangle \langle Q \rangle} B_{TQ}(\mathbf{p}) \quad (2.6)$$

where

$$B_T(\mathbf{p}) = \langle \delta T(\mathbf{r}) \delta T(\mathbf{r} + \mathbf{p}) \rangle \quad (2.7)$$

etc. Applying the Fourier transform to eq. (2.6) leads to

$$\varphi_n(\mathbf{k}) = \frac{A_T^2}{\langle T \rangle^2} \varphi_T(\mathbf{k}) + \frac{A_Q^2}{\langle Q \rangle^2} \varphi_Q(\mathbf{k}) + \frac{2A_T A_Q}{\langle T \rangle \langle Q \rangle} \varphi_{TQ}(\mathbf{k}) \quad (2.8)$$

where

$$f_n(\mathbf{p}) = \int_{R^3} \varphi_n(\mathbf{k}) \exp(i\mathbf{k} \cdot \mathbf{p}) d^3\mathbf{k} \quad (2.9)$$

For well developed (clear) atmospheric turbulence it can be shown that the spectral density function $\varphi_n(\mathbf{k})$ can be approximated with proper assumptions by [1,2]

$$\varphi_n(\mathbf{k}) = \frac{0.33 C_n^2 \exp(-\mathbf{k}^2/k_m^2)}{(\mathbf{k}^2 + 1/L_0)^{11/6}}$$

where L_0 is the largest turbulence scale (typically on the order of 100 meters) C_n^2 is known as the "index of refraction structure constant". It has to be estimated from the spectral density of the turbulent fluctuations in the atmosphere.

A long mathematical derivation based on the works of Tatarski and others [2,3] then yields that $C_n^2 \approx k_\theta/k_m$ where

$$k_\theta = - \frac{\langle \theta' v' \rangle}{\langle \text{grad } \bar{\theta} \rangle}$$

and K_M is the eddy diffusivity for momentum transfer which is related to the turbulent energy dissipation ϵ by

$$\epsilon = K_M (\text{grad } \bar{U})^2$$

Here $\bar{\theta}, \bar{U}$ are the average potential temperature \bar{U} the mean horizontal velocity and θ', v' represent the turbulent fluctuations.

To use this relations one must apply a proper algorithm to separate the mean part of the meteorological data from the turbulent fluctuations. The main objective of this

project was to use and compare different filtering methodologies in order to find the most appropriate for our purpose.

III. Data Filters.

The raw meteorological data regarding the wind velocity, pressure and temperature have to be decompose into the mean and fluctuating parts to derive a value for C_n^2 . This can be accomplished by different methods which lead in turn to different estimates for C_n^2 .

In the following we describe briefly the various techniques which were used in this project but due to space limitations we shall not be able to describe their technical aspects. To this end the reader is referred to the bibliography.

1. Time series approach [4]

In this method we fit a **local** low order polynomial to the data. The algorithm which is based on Box–Jenkins approach treat the meteorological data as due to a stochastic process. The decomposition, used properly, leads to residuals which have zero mean and variance. However the global ("smooth") structure of the mean quantities remains obscured by this process. One can refine this algorithm by the computation of the autocorrelation function for the data to determine the nature of the local filtering polynomial.

2. Low pass filter based on the FFT.[5]

In this approach we apply the fast Fourier transform to the data and by (inverse transform) remove from it the low frequency components. This leads to a smooth trigonometric approximation to the mean quantities which captures its global structure. However as the meteorological data contains some discontinuities the approximation fails at these points. As a refinement one can divide the data into "windows" of predetermined length (e.g. 2k to 10k) and apply the algorithm separately on each of these windows.

3. Least squares [5].

Here we fit a low order polynomial $p(x)$ to the data $\{d_i\}$ so that

$$S = \sum (p(x_i) - d_i)^2 \text{ is a minimum } (\ell^2 \text{ norm}) \quad (3.1)$$

As in the previous method one can refine the algorithm by applying it to data windows. In this way both the local and global structure of the mean quantities can be captured by the smoothing polynomial. Another possible refinement is to use an ℓ^p , norm $p > 1$ for the smoothing algorithm rather than ℓ^2 . E.g. in the ℓ^1 -norm eq. (3.1) will be replaced by

$$S = \sum |p(x_i) - d_i| \text{ is min.} \quad (3.2)$$

4. Spline smoothing filter [6].

In this algorithm one produces a natural cubic spline for a data set (t_i, x_i) with knots at all the t_i . However the smoothing spline does not interpolate for the x_i 's.

Instead it minimizes

$$\int^b S''(t)^2 dt$$

subject to the constraint

$$\sum \left| \frac{S(t_i) - x_i}{\omega_i} \right| < \sigma \quad (3.3)$$

where the ω_i 's are proper "data weights" and σ is a smoothing parameter. As a

refinement one can use various strategies to estimate the value of σ that is appropriate for a given data set.

5. Wavelet transform [7]

Wavelets represent a new development in applied mathematics which is used principally for signal processing. The basic idea is to use a function $\varphi(x)$ with compact support (called the mother wavelet) and apply to it a set of translations and dilatations to produce a basis of $L^2(\mathbb{R})$. This basis can be used then to expand some sets of functions "inexpensively" (i.e. with few terms). The hope is that in this manner one can capture both the local and global structure of the original signal.

In this project the idea was to use wavelets to capture the global structure of the meteorological signal and at the same time identify any local "jumps". To this end we used a wavelet transform package which was written by the "wavelet lab" at Yale University [8]. It turned out that the choice of $\varphi(x)$ was crucial for this transform to work. In fact the package contains four types of wavelets functions but only one detrended the data correctly, i.e. produced residuals with mean zero.

6. Convolution transform [9].

It is well known that

$$(\hat{f} * \hat{g})(s) = \hat{f}(s) \hat{g}(s)$$

where $*$ is the convolution operator and $\hat{}$ stands for the Fourier transform. It follows then that to eliminate from the spectrum of a signal $g(t)$ all frequencies greater than ω (i.e. consider the signal whose spectrum is $\hat{f}(s) \hat{g}(s)$ where

$$\hat{f}(s) = \begin{cases} 1 & 0 < s < \omega \\ 0 & s > \omega \end{cases},$$

we must compute the convolution

$$f * g = \int_{-\infty}^{\infty} f(x-y)g(x)dy$$

Similarly one can produce other filters on the spectrum of the original signal by appropriate choice of $\hat{f}(s)$.

This algorithm has been applied to the meteorological data but as it requires extensive computational resources we were unable to pursue this technique in depth as a paradigm for detrending the data.

IV Implementation

A software package was written to implement the various algorithms and strategies that were discussed in the previous section. The package is written in Fortran and uses the IMSL and TEKSIM libraries. It also uses a "wavelet lab. package" that was obtained from Yale University. Although the various parameters in these programs are currently set for the current data received from NASA, it will require a minor effort to adjust these as the need arises.

The following is a short description of the programs in this package:

plall.f – Read data from the original data file prepared by NASA and plot the data for each variable. As a result the relevant data can be identified and then extracted using condes.f

- condes.f – Extract the relevant data for the original NASA data file into a binary file this will speed the processing of the data in other programs.
- trend.f – Compute basic stat about the data and detrend it using differencing method. The order of the method should be set using the variable ndiff. The program also plots the detrended data.
- trend2.f – Same as above but plots the large component of the data after detrending or with minor modifications the original data and the large component superimposed on each other.
- trend3.f – Detrend data by the fft and plots the large component of after detrending.
- trend4.f – Detrend data by the least squares and plots the large component of the data after detrending
- trsp.f – Builds on trend.f and also compute the spectrum density for each of the six detrended variables in the file. To run one should set ndiff and iswver to determine the the stat window type.
- rdsp.f – Spectral analysis on the raw data. –similar to trsp.f
- psphl.f – Plots the spectral density from the ascii file produced by trsp.f or rdsp.f. half log plot
- pspll.f – Plots the spectral density from the ascii file produced by trsp/rdsp.f . log/log plot.

- `cfft.f` – Perform fft transform on the data using complex functions.
- `rfft.f` – Same as `cfft.f` but with real functions.
- `rfftb.f` – FFT and inverse to capture the large component of data with plots.
- `crall.f` – Take the binary file `mm*.bin` and converts it to six ascii files in preparation for use in the wavelet transform. Must run on unicon since binary files on other machines are not compatible.
- `wlsqsp.f` – Perform least squares fit to the data in a given window size. Then computes the residual and performs a spectral analysis on the residual of the whole data.
- `wfftsp.f` – Perform fft fit to the data in a given window size. Then computes the residual and performs a spectral analysis on the residual of the whole data.
- `wtrfsp.f` – Same as `wfftsp.f` but the spectral analysis is performed for each data window separately. the size of the window can be changed as a parameter.
- `wtrlsp.f` – Same as `wlsqsp.f` but the spectral analysis is performed for each data window separately.
- `conv02.f` – Detrend the data by the convolution method. the filter function used here corresponds to the normal distribution however the program is modular and this can be changed rather easily.

- conv04.f— Same as conv02 but with different parameters. In particular in this version the transform is being applied to a subset of the original data only and the computation proceeds on this subset only.
- cssp.f — Detrend data using spline approximation then computes the large and fluctuating parts and performs spectral analysis and plotting for each of the meteorological variables.
- cssp.f — Same as cssp but here the optimal smoothing parameter is computed by the algorithm.
- wcssp.f — Same as cssp.f but the smoothing is done in windows so that the local structure of the data can be taken into account.
- exw — A macro to execute the wavelet transform on the data. To use it one must apply crall.f to the file generated by condes.f.

In all the source code for this package occupies $\approx 200k$. A sample of the output from this program is given in the appendix.

V. Directions for Future Research.

The methods that were used in this project to detrend the meteorological data were based on the paradigm of "data smoothing" That is we assumed implicitly that the large component of the data is a smooth function which can be approximated in a window of reasonable size by a low order polynomial or some trigonometric function.

While this paradigm is based on a sound mathematical and scientific principle it is not generic. Thus it fails to establish an intrinsic relationship (based on some fundamental physics) between the data and the smoothing functions. Accordingly different choices of these functions will (and did) lead to different spectrum for the turbulent residuals. TO REMEDY THIS THE FILTER FUNCTION MUST BE "CUSTOMIZED" TO THE STRUCTURE OF THE METEOROLOGICAL DATA.

To accomplish this objective we would like to suggest at least three possible paradigms. These were not followed in this project so far due to their scope and the lack of time. To implement these ideas will require a preliminary stage of original research and development.

A. Karhunan-Loeve Basis functions.

This procedure was suggested by Lumely to treat meteorological data and isolate coherent structures in a flow. However to derive the functions that are needed for this analysis one has to solve an integral equation whose kernel is related to the raw data available about the flow. In this fashion the smoothing function carry an intrinsic relationship to the data that is being analyzed.

B. Intrinsic "Mother" wavelet.

The wavelets that were used in the current analysis were obtained from the "wavelet lab" at Yale University. The choice of the mother wavelets in this package have no intrinsic relationship to the meteorological data at hand. To improve this analysis the following problems have to be resolved.

1. Is it possible to define and use a "data fitted" mother wavelet.
2. To what extent the mother wavelet obtained in (1) is universal. i.e. is the same for different data sets.

3. What effect have different intrinsic mother wavelets on the spectrum of the turbulent residuals.
4. It it possible to relate the choice of the mother wavelet to flow invariants and characteristics (e.g vorticity and energy—dissipation).

C. Solution of Boussinesq eqs.

This is the fundamental approach whereby one solve the governing equations of the flow (with appropriate boundary conditions) to determine the spectral contents of the flow and it large scales. As the meteorological data is assumed to be a solution of these equations the results obtained can be used as a prototype for the detrending of the actual data. This approach will also lead to insights regarding the evolution of clear air turbulence and the factors the influence it.

Aside from these research directions we want to point out the following practical points regarding this project:

1. The software package that was developed use the graphic library "TEKSIM" which is available only at Philips lab. To make it portable for other location these calls should be replaced by others to widely available package of interactive graphics such as IMSL—EXP. Unfortunately this is not available at the lab at the present time.

2. The analysis of the data by the algorithms implemented hereby take approximately 100 cpu minutes on a scalar machine. If an adaption of this package for real time applications is desired the algorithm has to be refined so that results can be obtained in a span of few minutes.

3. In the current data set the sampling was done in intervals 0.2 seconds. This interval must be shortened considerably if the turbulent spectrum at higher frequencies is desired.

Bibliography

1. R.L. Fante – Wave propagation in Random Media, Prog. in optics 22 p. 341 (85).
2. Tatarski – Wave propagation in a turbulent medium. (McGraw Hill 1961).
3. E.M. Dewan – Optical Turbulence Forecasting AFGL–TR–80–0300 (1980)
4. G. Box and G.M. Jenkins – Time series analysis (Holden Day 1976).
5. C. de Boor – A practical guide to splines (Springer Verlag 1978)
6. C.F. Gerald – Applied Numerical Analysis (Addison Wesley 1983)
7. J. Morlet et al – Wave propagation and sampling theory, Geophysics 47 p. 202 (1982)
8. I. Daubechies – Ten Lectures on Wavelet (SIAM 1992)
R. Coifman et al – Wavelet analysis and signal analysis IEEE Trans Inform. The. 38 p. 713–718.

Appendix

STATISTICS OF ORIGINAL DATA

	psta	tsta	tas	u	v	w
mean	70.759	210.892	212.375	47.811	20.966	-0.877
variance	23.267	36.703	7.895	99.282	159.630	0.307
std. dev.	4.824	6.058	2.810	9.964	12.634	0.554
skewness	0.915	0.679	-0.336	-0.127	-0.275	0.364
kurtosis	0.103	-0.904	1.619	-1.160	-1.364	0.830
max	85.400	227.040	223.500	65.410	41.740	2.118
min	63.200	202.260	200.510	24.080	-3.000	-3.004
range	22.200	24.780	22.990	41.330	44.740	5.122
coef. var	0.068	0.029	0.013	0.208	0.603	-0.631
count	167199.	167199.	167199.	167199.	167199.	167199.

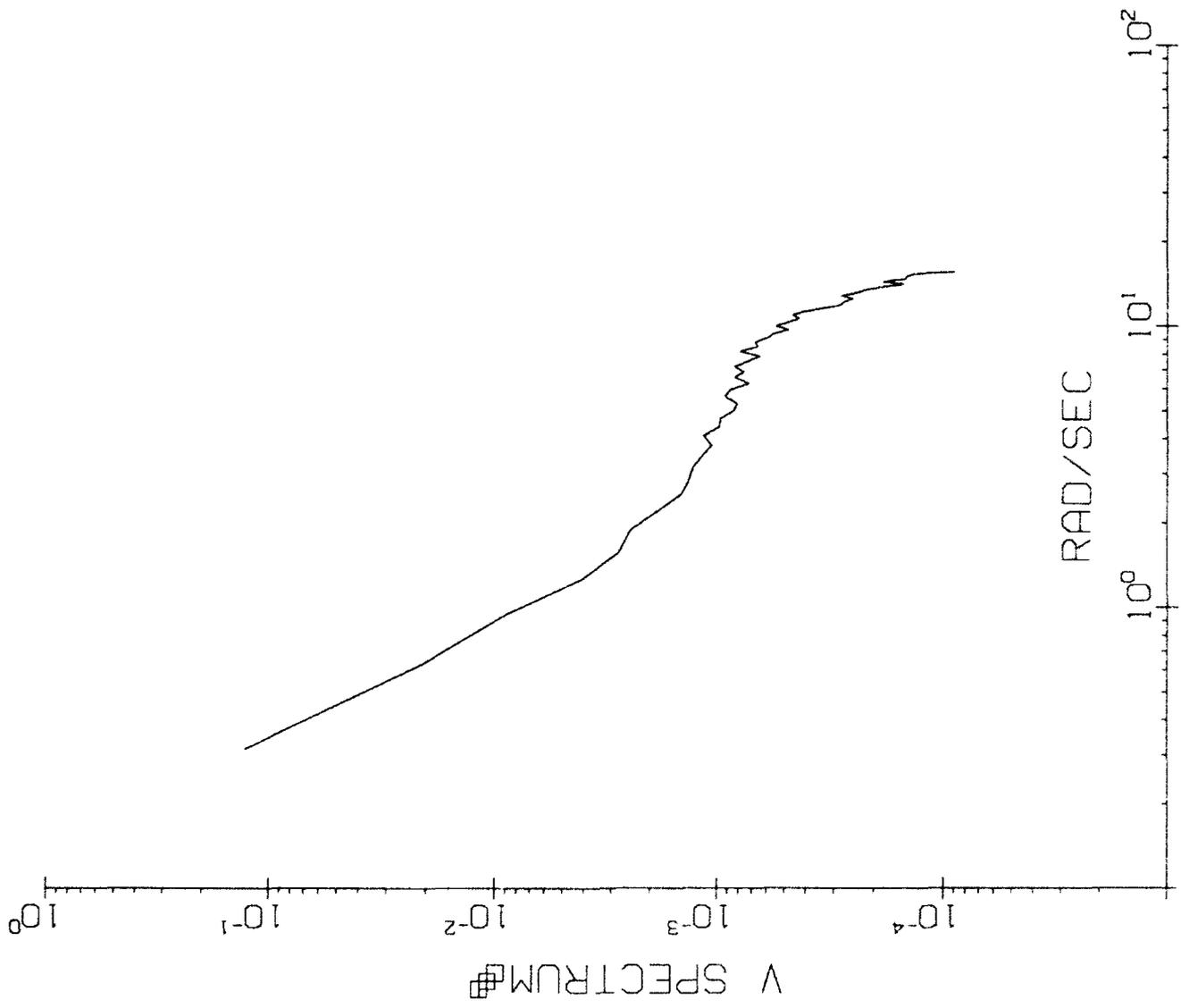
STATISTICS OF DETRENDED DATA
DIFFERENCING ORDER = 2

	psta	tsta	tas	u	v	w
mean	0.000	0.000	0.000	0.000	0.000	0.000
variance	0.022	0.019	0.090	0.074	0.033	0.009
std. dev.	0.150	0.137	0.300	0.272	0.182	0.097
skewness	0.010	-0.016	-0.014	-0.017	0.055	-0.058
kurtosis	0.981	0.093	0.451	0.635	3.617	6.178
max	0.800	0.640	1.880	1.910	2.500	1.228
min	-0.800	-0.700	-2.150	-2.280	-1.890	-1.715
range	1.600	1.340	4.030	4.190	4.390	2.943
count	167199.	167199.	167199.	167199.	167199.	167199.

1

SPECTRAL ANALYSIS FOR PSTA
using the window of Parzen

frequency	period	m=512	m=1024	m=2048
0.3132	20.06	0.1606E-06	0.1519E-06	0.1393E-06
0.6273	10.02	0.1060E-05	0.1064E-05	0.1107E-05
0.9415	6.67	0.3802E-05	0.3840E-05	0.3987E-05
1.2556	5.00	0.1055E-04	0.1065E-04	0.1110E-04
1.5698	4.00	0.2495E-04	0.2530E-04	0.2572E-04
1.8840	3.34	0.4489E-04	0.4340E-04	0.4036E-04
2.1981	2.86	0.8145E-04	0.7832E-04	0.7411E-04
2.5123	2.50	0.1334E-03	0.1310E-03	0.1253E-03
2.8264	2.22	0.1992E-03	0.1959E-03	0.1969E-03
3.1406	2.00	0.3031E-03	0.3169E-03	0.3277E-03
3.4548	1.82	0.3858E-03	0.3821E-03	0.3725E-03
3.7689	1.67	0.5640E-03	0.5695E-03	0.5559E-03
4.0831	1.54	0.6991E-03	0.6936E-03	0.6915E-03
4.3972	1.43	0.7672E-03	0.7684E-03	0.8028E-03
4.7114	1.33	0.9856E-03	0.9886E-03	0.1003E-02
5.0255	1.25	0.1164E-02	0.1156E-02	0.1227E-02
5.3397	1.18	0.1372E-02	0.1412E-02	0.1492E-02
5.6539	1.11	0.1537E-02	0.1555E-02	0.1539E-02
5.9680	1.05	0.1617E-02	0.1666E-02	0.1726E-02
6.2822	1.00	0.1939E-02	0.2037E-02	0.2136E-02
6.5963	0.95	0.1880E-02	0.1923E-02	0.1926E-02
6.9105	0.91	0.1970E-02	0.1978E-02	0.1986E-02
7.2247	0.87	0.2137E-02	0.2183E-02	0.2324E-02
7.5388	0.83	0.1918E-02	0.1888E-02	0.1861E-02
7.8530	0.80	0.2131E-02	0.2242E-02	0.2333E-02
8.1671	0.77	0.1893E-02	0.1882E-02	0.1878E-02
8.4813	0.74	0.1712E-02	0.1751E-02	0.1747E-02
8.7955	0.71	0.1874E-02	0.1950E-02	0.2007E-02
9.1096	0.69	0.1524E-02	0.1491E-02	0.1507E-02
9.4238	0.67	0.1461E-02	0.1503E-02	0.1602E-02



**A THERMAL MODELING OF DIRECTIONALLY WOVEN
CARBON-CARBON FIBER MATERIALS - PREDICTION OF
ORTHOTROPIC THERMAL CONDUCTIVITY**

**Amir Karimi
Associate Professor
Division of Engineering**

**The University of Texas at San Antonio
Division of Engineering
6900 N. Loop 1604 W.
San Antonio, Texas 78249-0665**

**Final Report for:
Summer Research Program
Phillips Laboratory**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.**

September 1992

A THERMAL MODELING OF DIRECTIONALLY WOVEN CARBON-CARBON FIBER MATERIALS - PREDICTION OF ORTHOTROPIC THERMAL CONDUCTIVITY

Amir Karimi
Associate Professor
Division of Engineering
The University of Texas at San Antonio

Abstract

This study investigates the thermal characteristics of directionally woven carbon-carbon (C-C) composite materials. It contains thermal modeling of composite materials for the prediction of orthotropic thermal conductivities. The prediction formula is based upon a thermal circuit analysis of a unit cell, describing the repetitive construction pattern within a directionally woven C-C composite material. It illustrates that a unit cell, itself, can be divided into smaller components, or sub-cells. A circuit approach is employed to formulate the effective directional thermal conductivities of a generic sub-cell, containing an arbitrarily oriented carbon fiber within the precursor matrix. The influence of the longitudinal and transverse thermal conductivities of carbon fiber, the thermal conductivity of the precursor matrix, and the contact resistance, are incorporated in the thermal modeling. The orthotropic thermal conductivity is then based upon thermal circuit analysis of the network of the sub-cells within a unit cell of a directionally woven C-C fiber composite material.

The model developed in this study may be used to predict the orthotropic thermal conductivities for several geometrical configurations of directionally woven carbon-carbon fibers including unidirectional (1-D), bi-directional (2-D), and multidirectional (3-D, and 4-D). The proposed model could be utilized for the examination and design of active and passive radiators for space applications. Existing thermal conductivity data necessary for verification of the model is very limited. Specifically, the transverse thermal conductivity data of carbon fibers is scarce. It is suggested that further investigations are required for thermal conductivity data collection and verification of the model. Creation of a data base on the thermal conductivities of individual carbon fibers and fabricated C-C composite materials is recommended for future studies. An extension for the application of this model is the prediction of effective thermal conductivity of wire screens used in heat pipes. The comparison of this model shows good agreement with available data.

A THERMAL MODELING OF DIRECTIONALLY WOVEN CARBON-CARBON FIBER MATERIALS - PREDICTION OF ORTHOTROPIC THERMAL CONDUCTIVITY

Amir Karimi

INTRODUCTION

Composite materials are used in a variety of industrial applications where extreme environments of heat and stresses exist including rocket motor nozzles, supersonic aircraft disc brakes, space vehicle structures, space radiators, etc. The ability to understand and to model the principal physical mechanism associated with the carbon-carbon (C-C or C/C) fiber materials is of considerable importance.

Thermal management of the electronic packages in modern spacecrafts requires utilization of advanced materials for heat dissipation into the space environment. Recent efforts have been directed towards the design, analysis, development and selection of light-weight, high conductivity composite materials. Carbon-carbon composite materials, in particular, have been considered as a part of the thermal management systems for spacecrafts [1-8]. These materials possess enhanced mechanical properties combined with thermal endurance, high temperature capability, excellent survivability in hostile environment, high specific strength and modulus, low coefficient of thermal expansion. New processing techniques has made it possible to form these materials into sheets and tubes. The potential payoffs for utilization of C-C composite materials include weight saving, ability to meet advanced survivability requirement, improved efficiency in radiator and advanced mirror designs, and structural material compatibility at extremely elevated temperatures.

Carbon-carbon composites are, in general, composed of a carbonaceous or graphitic matrix reinforced with carbon or graphite fibers. The carbon fibers in these materials are either dispersed, layered, or woven in certain orientation and configuration. In dispersed composite, the carbon fibers are randomly distributed in the precursor matrix. In layered composites, layers of carbon fibers or yarns are arranged in a preferred direction(s). In woven composites, carbon fibers are interlaced in an specific pattern within the precursor matrix. The woven carbon fibers in composites could have uni, bi-, or multidirectional alignments.

Manufacturing of C-C composite materials involve complicated multi-stage processes. Obviously, the processing technique depends largely on the application of finished material. Many attempts have been made to identify an optimum process for producing high thermal conductivity C-C composites. Commonly, fabrication C-C composite materials include [9]: 1) configurational arrangement of reinforcing phase; 2) stabilization of reinforcing grid, usually with

Phenolic resin; 3) matrix formation using successive cycles of impregnation, carbonization, and graphitization of high yield resins such as coal tar pitch.

Physical properties and transport properties of carbon fibers play a significant role in the design and analysis of C-C composite materials used in heat transfer systems and processes. The anisotropic characteristics of thermal conductivity typical of carbon fibers is useful in fabrication of composite materials for specific applications where favorable directional heat transfer is desired. For example, design of radiators or condenser fins for the thermal management of a spacecraft, where uniform temperature distribution is criteria, requires selection and utilization of a composite material having a planer-high-conductivity characteristic.

Selection of advanced composite materials for a specific thermal application requires knowledge of physical properties. In many situation it is necessary to design and fabricate a composite material that meets the requirements of an application. Manufacturing of C-C composites to test for thermal properties is too costly. Therefore, it is crucial to determine the heat transfer characteristics of these materials through analytical modeling. Design and optimization of C-C composite material require qualitative information including thermal conductivities of individual fibers and precursor matrix, orientation of carbon fibers, and fiber volume fraction distribution. The contact thermal resistance at the interface of carbon fiber and the precursor matrix may also influence the effective thermal behavior.

OBJECTIVES

The prediction of the effective thermal conductivity and the major uncertainties associated with it play a vital role in modelling, design and fabrication of carbon-carbon composite materials. The main objective of this effort is to develop an analytical model for prediction of directional thermal conductivity of woven carbon-carbon materials to support USAF Phillips Laboratories current mission of analysis and design of radiators for satellite applications. The specific goals of this work are to identify those parameters that contribute to the effective thermal conductivity, to develop prediction formulation for thermal conductivity, and make a comparative study of the model with the available data. These will require a detailed analysis of anisotropic heat transfer mechanism in composite materials. This effort would be specifically directed towards the development of prediction formulas of thermal conductivity for multidirectional woven carbon fibers.

This study will be limited to examination of conduction mechanism in the directionally woven C-C material. Specifically, the modelling and formulation of thermal conductivities for various configuration of the woven carbon fibers will be the focus of this research. The result of this study could be applied to other composite materials of similar configuration. In this

investigation, the modeling of the thermal conductivity for C-C composite materials is based on a simple thermal circuit model well the thermal resistance of each constituent in the composite material are contributing to the overall thermal conductivity. A more detailed and complete analysis requires to verify the validity of the proposed model.

ANALYSIS

Composite materials can be classified into three broad categories [10]: layered, fibrous, and dispersed. The focus of this study we will be on the special cases of the first two classifications, mainly, layered carbon-carbon fibers and multi directional woven fibers.

In this work we will attempt to formulate the effective thermal conductivity of directionally woven carbon carbon fiber materials. Formulation of the problem will include the longitudinal and transverse thermal conductivities of carbon fibers. It also includes the thermal conductivity of the precursor matrix. Since studies have shown that the transverse thermal conductivity of carbon fibers are in general one or two order of magnitude smaller than the longitudinal thermal conductivity it might be possible to ignore the contributions made by transverse thermal conductivity of carbon fibers.

In directionally woven C-C composite materials, the in-plane fibers are referred to as warp fibers (or stuffers) and fill fibers (or pickers) . The directions within the plane weave are referred to as warp (in 0 direction) and fill (in 90 direction) direction. The out of plane fibers are called web web (weaver interlock). The web angle is the angle between web yarns and the planer surface. Sometimes, instead of individual fibers, yarns are used in the construction of reinforced carbon fiber matrix. Yarn consists of a bundle (as high as thousands) of carbon fibers. In real cases when yarns are used in the C-C reinforcement matrix the shapes are not circular cylinders, they are deformed elliptical or oval shape.

In unidirectional carbon carbon composites where all carbon fibers are oriented in a single direction the principal thermal conductivity will be in the direction of fiber alignment. In a bi-directional C-C composite the carbon reinforcement fibers intersect in a planer surface with an angle. In most 2-D C-C composite materials the fibers in the fill direction are normal to the fibers in the fibers in the warp direction. Schematic models of 2-D woven carbon-carbon composite materials are shown in Figures 1 through 3. The fibers in the fill and warp directions could be straight or they might encounter interweaving which results in an angle of interweaving or crimp angle. Obviously, the principle conductivity would be in the the planer surface (fill and warp directions) and the minor conductivity would be in out -of-plane direction (normal to the planer surface).

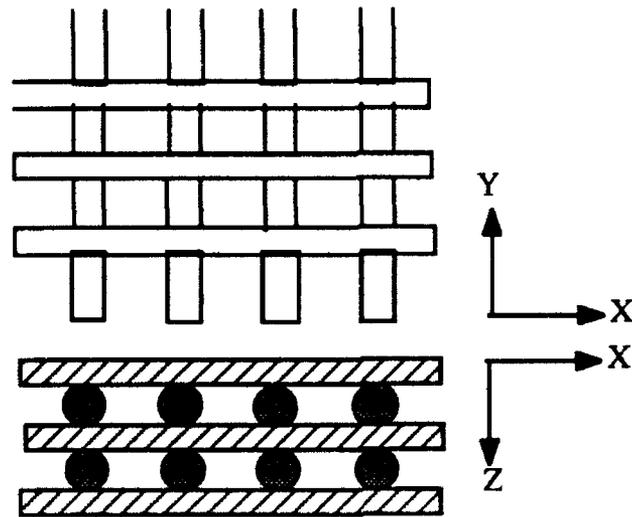


Fig. 1. Schematic representation of a 2-D C-C fiber material -- no interweaving

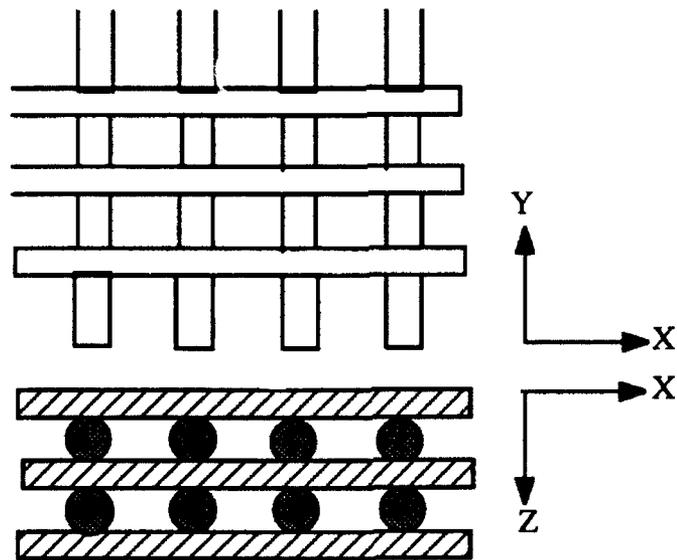


Fig. 2. Schematic representation of a 2-D C-C fiber material -- interweaving in Y direction

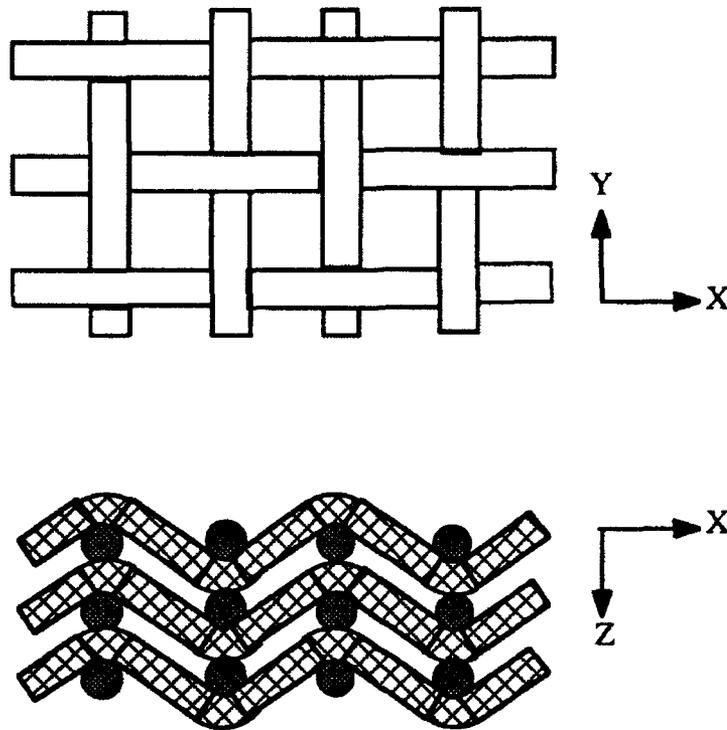


Fig. 3. Schematic representation of a 2-D C-C fiber material -- interweaving in both x and y directions

In 2-D C-C composites the thermal conductivities in the fill and warp direction would be the same if the same type of fibers are used in equal spacing in both directions. In 3-D directional orthogonal weaves the proper volume fraction of carbon fibers in the principal axis direction of composite, could provide desirable in-plane and out-of-plane thermal conductivity. In 3-D and 4D weave angle interlock he angle interlock weaves can provide a tailorable adequate balance of in-plane and out-of-plane mechanical and thermophysical properties.

Modeling of woven carbon fibers are done via a unit cell approach. The unit cell of a woven fabric describes the geometry of the fabric. Parameters influencing the model may include weaving angle, fiber types, yarn counts, reinforcement volume fraction, repeat pattern. We propose to use a thermal resistance analogy model to study the conduction process in the C-C composite materials. This requires defining a unit cell based on the reappearing pattern of the constituents within the composite material.

Formulation of Problem:

Unit cells in directionally woven materials can be divided into a number of sub-cells. For example, a unit cell identifying the repeating pattern of the constituents within a 2-D

interweave composite material consists of four basic compartments, as shown in Fig. 4. In compartments 1 through 3 each sub cell consists of a carbon fiber within the precursor matrix, as shown in Fig.5.. Section 1 shows two sub-cells where the warp and fill fibers cross each other; sections 2 and 3 represent parts of warp and fill fibers, respectively; and section 4 denotes the precursor matrix. To determine the effective directional thermal conductivities of the unit cell, one must first determine the effective directional thermal conductivities of each sub-cell and use the results in a thermal resistance network to evaluate the directional thermal resistance of the unit cell.

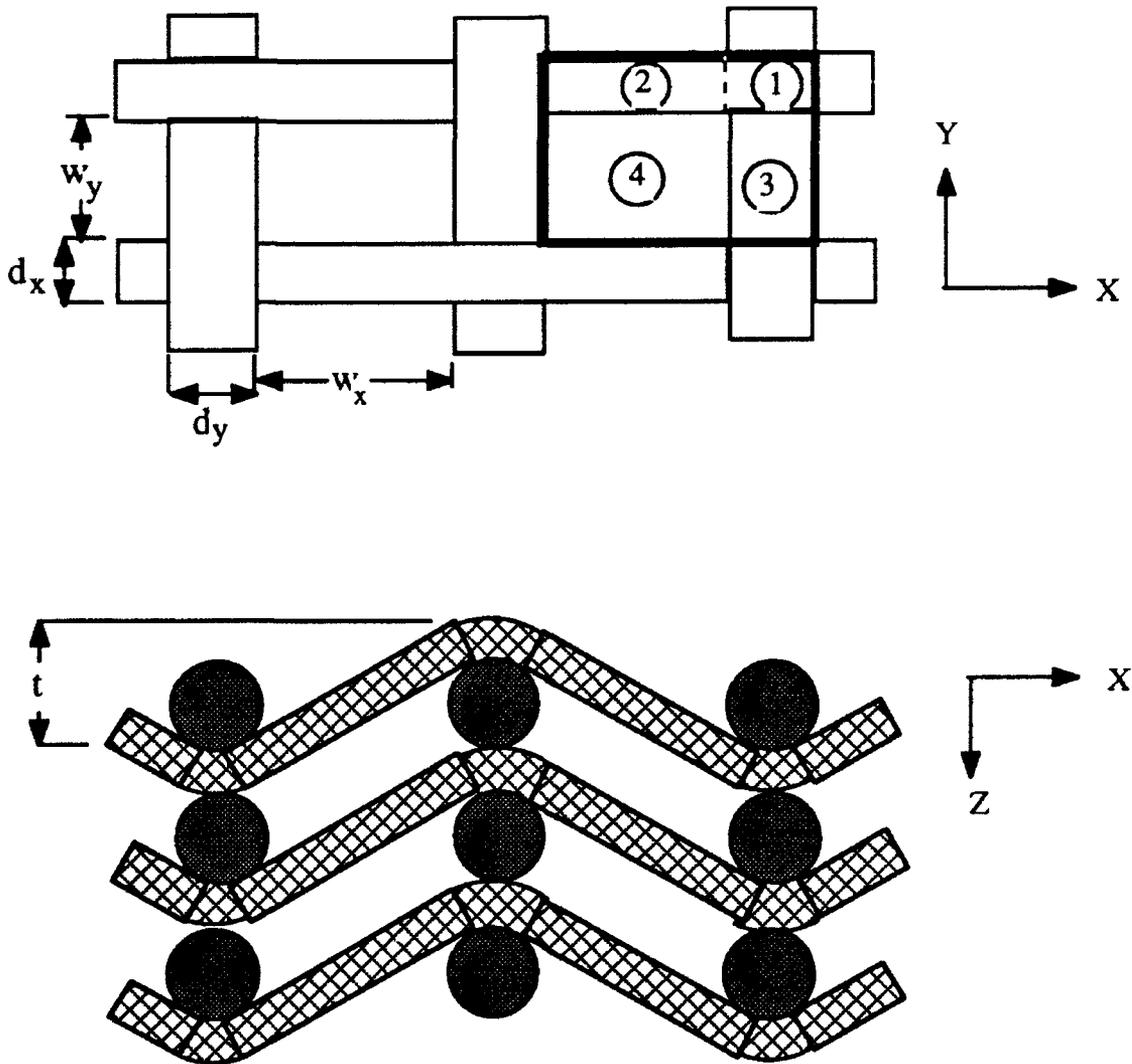


Fig. 4. Schematic representation of the sub-cells within a 2-D C-C fiber material

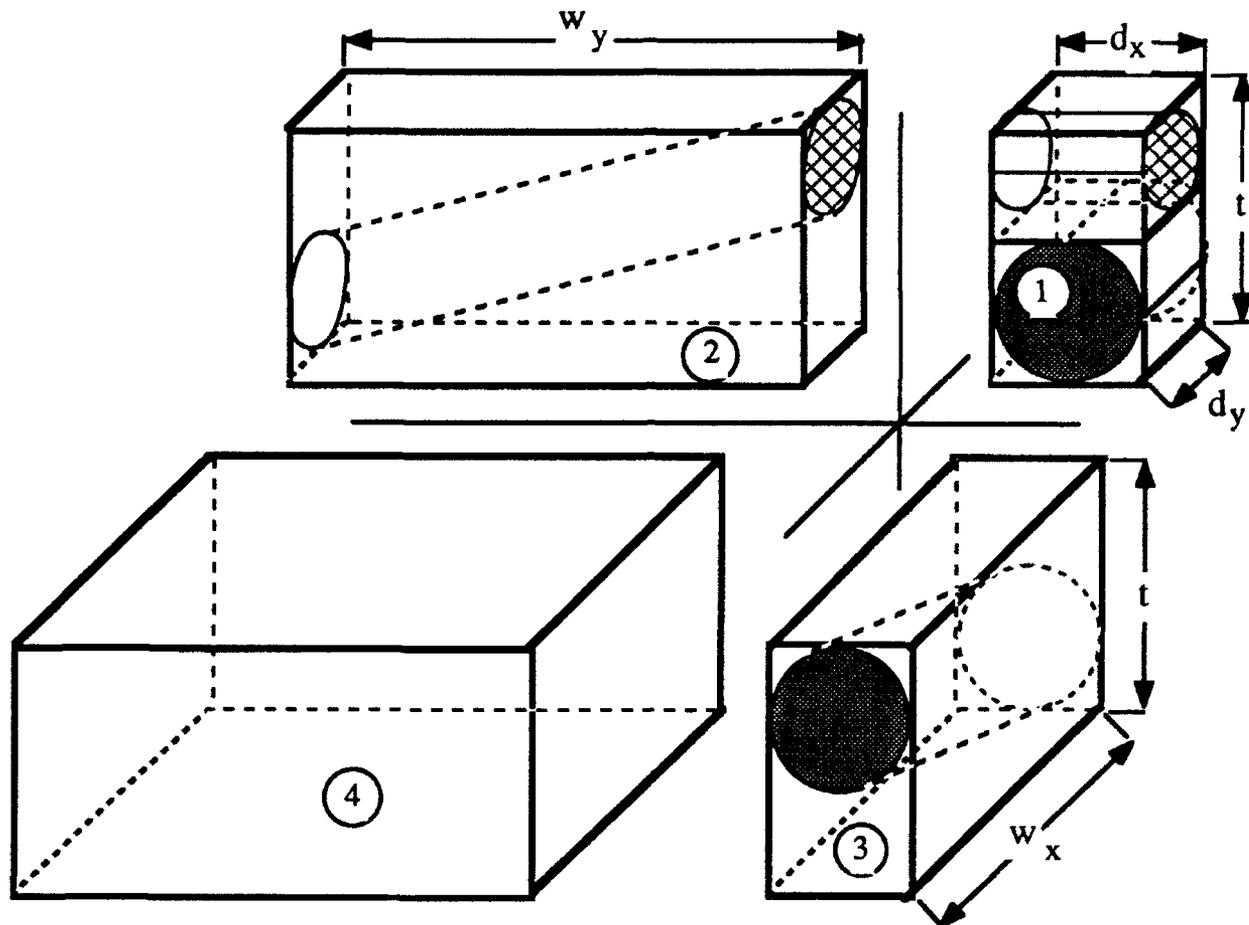


Fig. 5. Schematic representation of individual sub-cells within a 2-D C-C fiber material

The assumptions made in modeling the thermal conductivity of woven C-C composite material are:

1. carbon fibers are in circular cylindrical shapes;
2. fibers are orthotropic (the thermal conductivity axis are in the principal axis);
3. major or principal thermal conductivity of fiber is in the axial direction;
4. minor thermal conductivity of the fiber in the radial direction;
5. the porosity is completely filled with bonding materials (precursor matrix);
6. thermal conductivity of the precursor matrix is isotropic (uniform thermal conductivity in all directions);
7. the effect of curvature on thermal conductivity resulting from fiber interweave or interlock is ignored

Examination of sub-cells in various types of directionally woven materials, reveals that the fiber orientation within a sub-cell can be represented in a general configuration, as shown in Fig. 6. In this model the generic sub-cell is represented in Cartesian coordinates containing a single fiber within the precursor matrix. The sub-cell has dimensions l_x , l_y , and l_z in x, y, and z directions, respectively. The carbon fiber is represented by a circular cylinder rotated in x-y plane making an angle of ϕ with x axis and then rotated in a direction normal to x-y plane with an angle θ , as shown in Fig. 7. We note that the carbon fiber is sectioned by x-y plane in an elliptical shape.

For convenience of analysis, we suggest to transform the inclined circular cylinder fiber, in the generic sub cell, into a rectangular prallelepiped of equivalent volume, as shown in Fig. 8. It includes the transformation of the directional thermal conductivities of the carbon fibers into components in the principal axis directions, x,y, and z. The transformation can be described in two step. First the circular cylindrical fiber is transformed into an inclined square bar of equivalent cross section (equivalent volume). The side of square, a, is related to the diameter of fiber, d, by:

$$a = \sqrt{\pi} d / 2 = 0.8862 d \quad (1)$$

In this step of transformation it is assumed that the longitudinal and the transverse thermal conductivities of the square fiber is the same as those for the circular fiber. Note that in this stage of transformation the projection of the square fiber's cross-section onto the y-z plane yields a rectangle of dimensions b, and c. The second step in is to transform the square fiber to a rectangular prallelepiped of equivalent volume parallel to the principal axis, as shown in Fig. 8. The dimensions of the rectangle, b and c, are related to the diameter of fiber by:

$$b = a / \cos \phi = \sqrt{\pi} d / (2 \cos \phi) = 0.8862 d / \cos \phi \quad (2)$$

and

$$c = a / \cos \theta = \sqrt{\pi} d / (2 \cos \theta) = 0.8862 d / \cos \theta \quad (3)$$

respectively.

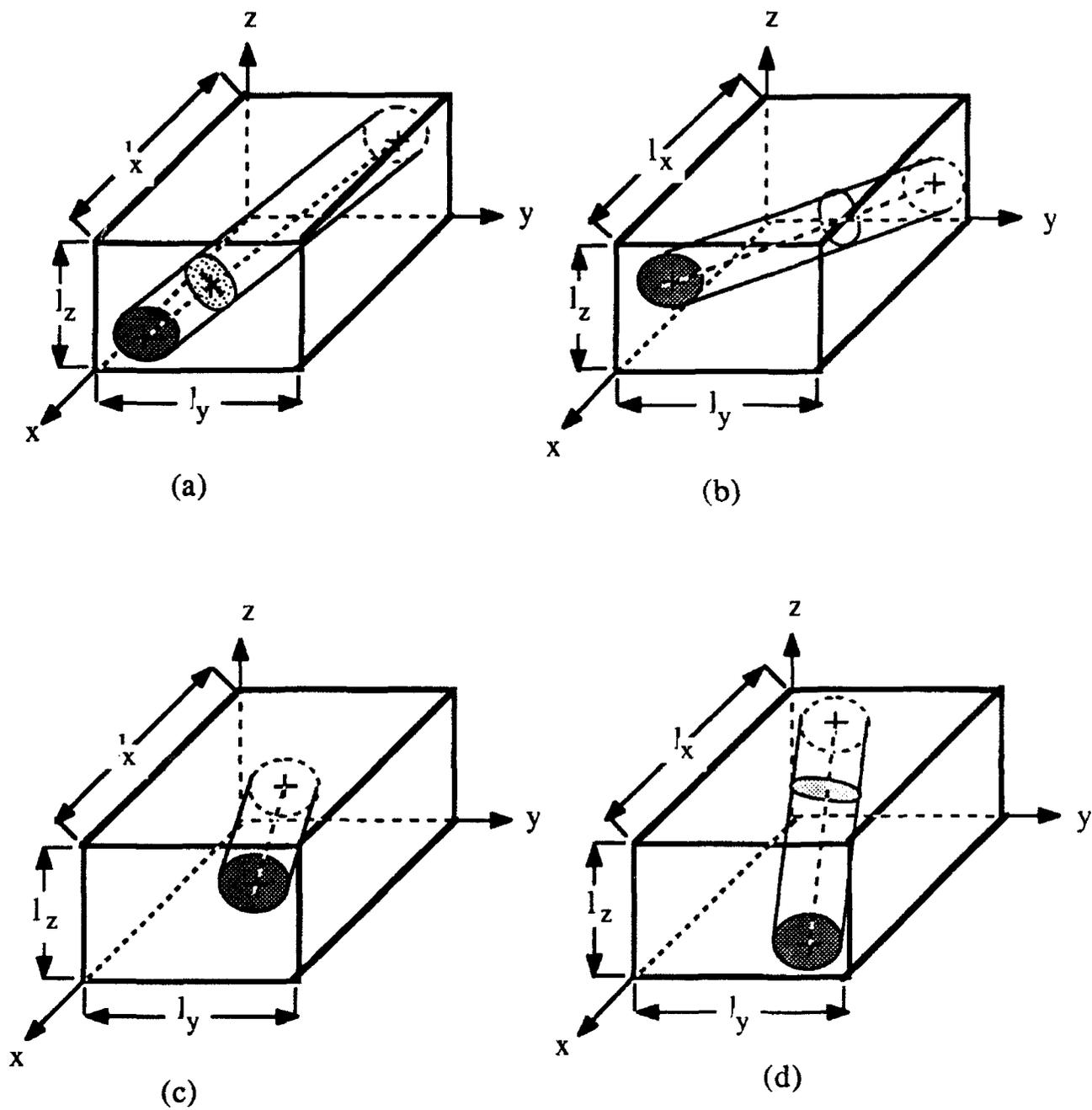


Fig.6. General representation of a carbon fiber in precursor matrix

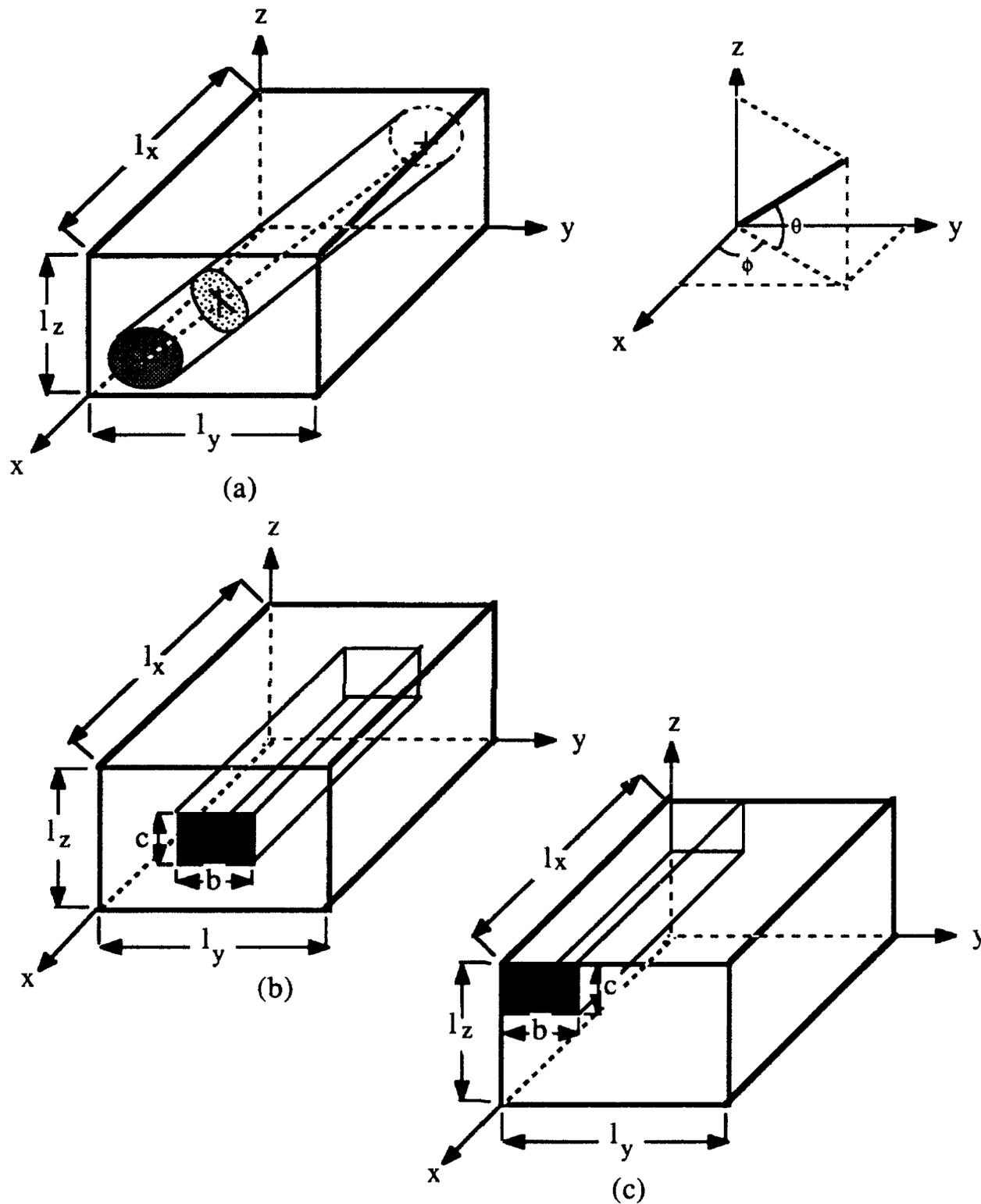


Fig. 8. Transformation of carbon fiber from a circular cylinder to rectangular bar bonding matrix

Resolution of the longitudinal and transverse thermal conductivities of the original carbon fiber onto the principal axis yield the orthotropic thermal conductivities for the transformed fiber. The directional thermal conductivities can be presented as:

$$\kappa_x = k_x/k = \cos^2\theta \cos^2\phi + \beta (\sin^2\phi \cos^2\theta + \sin^2\theta \cos^2\phi) \quad (4)$$

$$\kappa_y = k_y/k = \cos^2\theta \sin^2\phi + \beta (\cos^2\phi + \sin^2\theta \sin^2\phi) \quad (5)$$

$$\kappa_z = k_z/k = \sin^2\theta + \beta (\cos^2\theta + \sin^2\theta \sin^2\phi) \quad (6)$$

where κ_x , κ_y , and κ_z are the dimensionless orthotropic thermal conductivities; k_x , k_y , and k_z are the thermal conductivities of the transformed carbon fiber in x, y, and z directions respectively; $\beta = k'/k$; k and k' denote the longitudinal and transverse thermal conductivities of the original carbon fiber.

Effective Directional Thermal Conductivity of Sub-cell

Effective directional thermal conductivities of the sub-cell could be determined via circuit analysis of the network of thermal resistances within the sub-cell. We further notice that the networks of thermal resistances in the sub-cells, represented in Figs. 8-b and 8-c- are equivalent. In the x direction The effective thermal resistance of the generic sub-cell in the x direction is the result of two resistances connected in parallel,

$$R_x = \frac{R_{1,x} R_{2,x}}{R_{1,x} + R_{2,x}} \quad (7)$$

where

$$R_{1,x} = \frac{l_x}{k_x (b c)} \quad (8)$$

and

$$R_{2,x} = \frac{l_x}{k'' (l_y l_z - b c)} \quad (9)$$

$R_{1,x}$ and $R_{2,x}$ are the effective thermal resistances of the transformed carbon fiber and the precursor matrix, respectively, in the x direction; and k'' denotes the thermal conductivity of the precursor matrix.

Substituting equations (3) through (5), (8) and (9) into equation (7) the effective thermal resistance of sub-cell in the x direction can be expressed as:

$$R_x = \frac{l_x}{k (l_y l_z) [\gamma + (\kappa_x - \gamma) (\delta_y \delta_z)]} = \frac{l_x}{k_{e,x} (l_y l_z)} \quad (10)$$

where, $k_{e,x}$ is the effective thermal conductivity in the x direction, $\gamma = k''/k$, $\delta_y = b/l_y$, and $\delta_z = c/l_z$. Therefore, the effective thermal conductivity of the sub-cell in the x direction can be expressed as:

$$k_{e,x} = k[\gamma + (\kappa_x - \gamma) (\delta_y \delta_z)] \quad (11)$$

The network of thermal resistances of the sub-cell in y direction include an additional resistor to account for the effect of contact resistance and the influence of processing on thermal conductivity. It consists of four thermal resistors. The thermal resistance of the transformed fiber is given as:

$$R_{1,y} = \frac{b}{k_y (c l_x)} ; \quad (12)$$

The thermal resistance resulting from the contact resistance and the influence of processing is represented by:

$$R_{c,y} = \frac{\alpha}{c l_x} \quad (13)$$

where α is the contact resistance, in $m^2 \cdot ^\circ C / W$. Note that $R_{1,y}$ and $R_{c,y}$ are in series and the resulting thermal resistance is in parallel with the thermal resistance for the region below the fiber (region defined by $b \times (l_z - c)$), $R_{2,y}$, given by:

$$R_{2,y} = \frac{b}{k'' l_x (l_z - c)} ; \quad (14)$$

The thermal resistance for the region defined by $l_z \times (l_y - b)$, $R_{3,y}$, expressed by

$$R_{3,y} = \frac{(l_y - b)}{k'' (l_x l_z)} ; \quad (15)$$

The thermal resistance resulting from the combination of $R_{1,y}$, $R_{2,y}$, and $R_{c,y}$, is in series with $R_{3,y}$. Therefore, the effective thermal resistance in y direction may be expressed

$$R_y = \frac{(R_{1,y} + R_{c,y}) R_{2,y}}{(R_{1,y} + R_{c,y}) + R_{2,y}} + R_{3,y} \quad (16)$$

Combining equations (3) through (5) and (12) through (16) the following expression yields for the effective thermal resistance in y direction:

$$R_y = \frac{l_y}{k_{e,y} (l_x l_z)} = \left[\frac{l_y}{k (l_x l_z)} \right] \left\{ \frac{\gamma \delta_y (\delta_y + \epsilon_y \kappa_y) + (1 - \delta_y) [\gamma (1 - \delta_z) (\delta_y + \epsilon_y \kappa_y) + \delta_y \delta_z \kappa_y]}{\gamma [\gamma (1 - \delta_z) (\delta_y + \epsilon_y \kappa_y) + \delta_y \delta_z \kappa_y]} \right\} \quad (17)$$

where, $k_{e,y}$ is the effective thermal conductivity in the y direction, and $\epsilon_y = \alpha k / l_y$. Thus, the effective thermal conductivity of the sub-cell in the y direction can be expressed as:

$$k_{e,y} = k \left\{ \frac{\gamma[\gamma(1 - \delta_z)(\delta_y + \epsilon_y \kappa_y) + \delta_y \delta_z \kappa_y]}{\gamma \delta_y (\delta_y + \epsilon_y \kappa_y) + (1 - \delta_y)[\gamma(1 - \delta_z)(\delta_y + \epsilon_y \kappa_y) + \delta_y \delta_z \kappa_y]} \right\} \quad (18)$$

Similarly, the network of thermal resistances of the sub-cell in z consists of four thermal resistors. The thermal resistance of the transformed fiber is given as:

$$R_{1,z} = \frac{c}{k_z (b l_x)} \quad (19)$$

The thermal resistance resulting from the contact resistance and the influence of processing is represented by:

$$R_{c,z} = \frac{\alpha}{b l_x} \quad (20)$$

Again, $R_{1,z}$ and $R_{c,z}$ are in series and the resulting thermal resistance is in parallel with the thermal resistance for the region defined by $c \times (l_y - b)$, $R_{2,z}$, given by:

$$R_{2,z} = \frac{c}{k'' l_x (l_y - b)} \quad (21)$$

The thermal resistance for the region defined by $l_y \times (l_z - c)$, $R_{3,y}$, expressed by

$$R_{3,z} = \frac{(l_z - c)}{k'' (l_x l_z)} \quad (22)$$

The combination of $R_{1,z}$, $R_{2,z}$, and $R_{c,z}$, is in series with $R_{3,z}$. Therefore, the effective thermal resistance in z direction may be expressed

$$R_z = \frac{(R_{1,z} + R_{c,z}) R_{2,z}}{(R_{1,z} + R_{c,z}) + R_{2,z}} + R_{3,z} \quad (23)$$

Combining equations (3) through (5) and (19) through (23) the following expression yields for the effective thermal resistance in y direction:

$$R_z = \frac{l_z}{k_{e,z} (l_x l_y)} = \left[\frac{l_z}{k (l_x l_y)} \right] \left\{ \frac{\gamma \delta_z (\delta_z + \epsilon_z \kappa_z) + (1 - \delta_z)[\gamma(1 - \delta_y)(\delta_z + \epsilon_z \kappa_z) + \delta_y \delta_z \kappa_z]}{\gamma[\gamma(1 - \delta_y)(\delta_z + \epsilon_z \kappa_z) + \delta_y \delta_z \kappa_y]} \right\} \quad (24)$$

where, $k_{e,z}$ is the effective thermal conductivity in the z direction, and $\epsilon_z = \alpha k / l_z$. Thus, the effective thermal conductivity of the sub-cell in the z direction can be expressed as:

$$k_{e,z} = k \left\{ \frac{\gamma[\gamma(1 - \delta_y)(\delta_z + \epsilon_z \kappa_z) + \delta_y \delta_z \kappa_y]}{\gamma \delta_z (\delta_z + \epsilon_z \kappa_z) + (1 - \delta_z)[\gamma(1 - \delta_y)(\delta_z + \epsilon_z \kappa_z) + \delta_y \delta_z \kappa_z]} \right\}$$

where, $k_{e,y}$ is the effective thermal conductivity in the y direction, and $\epsilon_y = \alpha k / l_y$. Thus, the effective thermal conductivity of the sub-cell in the y direction can be expressed as:

Procedure:

The procedure for development of prediction formulas for thermal conductivity is as follows:

1. identifying a repeating unit cell for the woven C-C fiber composite material;

2. dividing a unit cell into smaller sub-components;
3. transforming the circular carbon fibers into Cartesian coordinates;
4. developing the thermal resistance formulation for each component of the sub-cell, using the directional thermal conductivities of the fiber element and the isotropic values for the thermal conductivity of bonding materials (precursor matrix);
5. determining the directional effective thermal resistance of each sub-cell using an electrical circuit analogy for the sub cell, (the components of each sub cell make up a network of serial and parallel thermal resistances);
6. determining the effective directional thermal conductivities of each unit cell using the directional thermal conductivities of sub-cells(again the sub-cell within each cell will make up a network of serial and parallel thermal resistances);
7. the results should provide effective thermal conductivities in longitudinal and transverse direction of fabric materials.

DISCUSSION

Thermal modeling of C-C composite materials requires a data base of the existing thermal conductivity data on various type of fibers and filler matrix in order to evaluate the contribution of each constituents to the directional thermal conductivity of finished product. Verification of the model also requires the knowledge of directional thermal conductivities of the fabricated material.

Experimental data for thermophysical and transport properties of carbon fibers is very limited, and test data for C-C composite materials are sparse. Our understanding of many important physical phenomena occurring in the processing of C-C composite materials remains incomplete. Commonly, the analysis of heat transfer process relies on experimentation. The approach adopted by the majority of the investigators has been an experimental approach for determination of thermal conductivity of fabricated material.

One of the main problems in analyzing the heat transfer phenomenon in carbon-carbon materials is the lack of unified thermophysical property test data base. It is extremely difficult to measure thermal conductivity of individual fibers. There are few experimental data available for the longitudinal thermal conductivity of individual carbon fibers [11-19]. The transverse thermal conductivity data is even more scarce. We have not been able to find any data for the transverse thermal properties of fibers. Therefore in modeling these materials one must resort to theoretical data for transverse thermal conductivity.

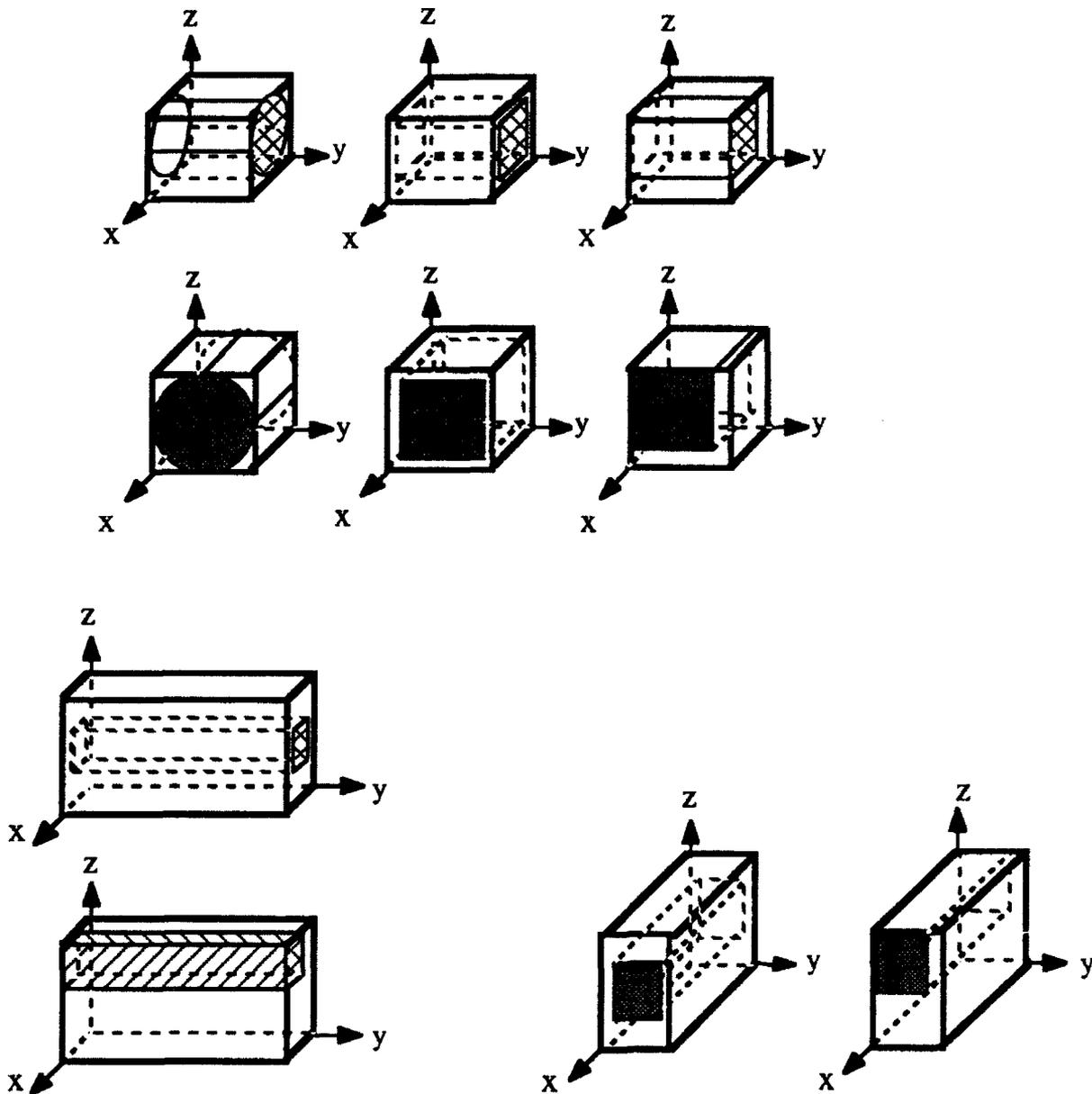


Fig. 9. Transformation in a 2-D woven C-C composite material

The thermal conductivity of C-C composite materials might be highly sensitive to the processing procedures and the heat treatment used in manufacturing [20-22]. Effect of precursor matrix (effect of processing parameters, density effect, thickness effect) might be an important contributing factor to the effective thermal conductivity.

The prediction formulas for thermal conductivity, developed in this work, may be applied to a variety of composite materials. Figure 9. shows the transformation of fibers in the sub-cell of a woven 2-D composite materials. Ignoring, the transverse thermal conductivities of the carbon

fibers, we used this configuration to compare the model with the existing data. The model gives predictions for the directional thermal conductivities, within 20% of the existing experimental data. One must note that there is a lack of agreement on experimental data for C-C composite materials. The model was also used for a comparison with Chang model [23] for the effective thermal conductivity of wire screens used in heat pipe application. Our model showed an improved prediction of thermal conductivity.

CONCLUSIONS AND RECOMMENDATIONS

Thermal characteristics of directionally woven C-C composite materials is discussed. A unit cell model is employed to describe the repetitive construction pattern within directionally woven C-C composite materials. It is shown that a unit cell can be divided into smaller components or sub-cell. A circuit approach is employed to formulate the effective directional thermal conductivities of a generic sub-cell, containing an arbitrarily oriented carbon fiber within the precursor matrix. The influence of the longitudinal and transverse thermal conductivities of carbon fiber, the thermal conductivity of the precursor matrix, and the contact resistance, are incorporated in the thermal modeling. The orthotropic thermal conductivity is then based upon thermal circuit analysis of the network of the sub-cells within a unit cell of a directionally woven C-C fiber composite material.

The model developed in this study may be used to predict the orthotropic thermal conductivities for several geometrical configurations of directionally woven carbon-carbon fibers including unidirectional (1-D), bi-directional (2-D), and multidirectional (3-D, and 4-D). The proposed model could be utilized for the examination and design of active and passive radiators for space applications. Existing thermal conductivity data necessary for verification of the model is very limited. Specifically, the transverse thermal conductivity data of carbon fibers is scarce. It is suggested that further investigations are required for thermal conductivity data collection and verification of the model. Creation of a data base on the thermal conductivities of individual carbon fibers and fabricated C-C composite materials is recommend for future studies. An extension for the application of this model is the prediction of effective thermal conductivity of wire screens used in heat pipes. The comparison of this model shows good agreement with available data .

NOMENCLATURE

- | | |
|---|---|
| a | $\sqrt{\pi} d / 2 = 0.8862 d$ |
| b | $a / \cos \phi = \sqrt{\pi} d / (2 \cos \phi) = 0.8862 d / \cos \phi$ |
| c | $a / \cos \theta = \sqrt{\pi} d / (2 \cos \theta) = 0.8862 d / \cos \theta$ |

d	diameter of carbon fiber
c_p	specific heat
k	axial thermal conductivity of carbon fiber
$k_x, k_y, \text{ and } k_z$	effective thermal conductivities of a unit cell sub-component in x, y, and z directions, respectively
k''	thermal conductivity of bonding material
k'	transverse conductivity of carbon fiber
$l_x, l_y, \text{ and } l_z$	dimensions of a unit cell sub-component in x, y, and z directions, respectively
R	thermal resistance, $^{\circ}\text{C}/\text{W}$
α	contact resistance, $(\text{m}^2 \cdot ^{\circ}\text{C}/\text{W})$
β	k'/k
δ_y	b/l_y
δ_z	b/l_z
ϵ_y	$\alpha k / l_y$
ϵ_z	$\alpha k / l_z$
γ	k''/k
$\kappa_x, \kappa_y, \text{ and } \kappa_z$	$k_x / k, k_y / k, \text{ and } k_z / k, \text{ respectively}$
$l_x, l_y, \text{ and } l$	unit cell-sub-component length, width, and height, respectively
ρ	density
ϕ	rotational angle of carbon fiber in the x-y plane
θ	rotational angle of carbon angle in a plane normal to the x-y plane

REFERENCES

- [1] "Carbon-Carbon Primary Structure," Interim Report , Boeing Aerospace and Electronics Aerospace Systems Technologies, WFWAL-TR-89-3007, 1987
- [2] "Carbon-Carbon Radiator for Space Applications, " Vol. 1, Revised, Technical Proposal, , Martin Marietta, , Astronautics Group, January 1992.
- [3] Engle, G.B., Tallon, J.A., and Graves, R.A., "High Thermal Conductivity, High Strength Carbon-Carbon Composites," Symposium on Advanced Materials, 1990??, PP. 319-325.
- [4] "Ford Carbon/Carbon Project Proposal," Airflow Sciences Corporation, July 1990.
- [5] Rawal, S., Misra, M.S., and Scheimberg, H.R., "Carbon-Carbon Radiator for Spacecraft Applications," Final Report , Martin Marietta, Materials Technology, Project D-81-R, 92-D-81-R-001, March 1992.
- [6] Rubin, L., "Characterization of High-Conductivity 3D Braid Carbon-Carbon," Aerospace Report No. ATR-89(4762)-1, Materials Science Laboratory, The Aerospace Corporation, El Segundo, CA, 1988.
- [7] "Space Carbon-Carbon Composites," Interim Report , Lockheed Missile & Space Co., WRDC-TR-89-4143, 1989.
- [8] Sullivan, B.J., and Rosen B.W., "Microstructural Evaluation of Carbon-Carbon Constituent, Bundle and Multi-Directional Composite Properties," Symposium on Advanced Materials, 1990, PP. 327-343.
- [9] Dhami, T.L., Manocha, L.M., and Bahl, O.P., "Oxidation Behavior of Pitch Based Carbon Fibers," Carbon, Vol. 29, No. 1, 1991, pp. 51-60.
- [10] Peterson, G.P., and Fletcher, L.S., "On the Thermal Conductivity of Dispersed Ceramics," ASME Journal of Heat Transfer, Vol. 111, 1989, pp. 824-829.
- [11] Cranch G.E., "Unique Properties of Flexible Carbon Fibers," Proceedings of fifth Conference on Carbon, Vol. I, The Macmillan Company, 1962, pp 589-594.
- [12] Deshpande, M.S., and Bogaard, R.H., "Evaluation of Specific Heat Data for POCO Graphite and Carbon-Carbon Composites," Pelnum Press, 1985, pp. 45-54.
- [13] Hagio, T., Nakamizo, M., and Kobayashi, K., "Thermal Conductivities and Raman Spectra of Boron-Doped Carbon Materials," Carbon, Vol. 25, No. 5, 1987, pp. 637-639.
- [14] Nightingale, R.E., Yoshikawa, H.H., and Losty, "Chapter Six - Physical Properties," Nuclear Graphite, Edited by R.E Nightingale, Academic Press, New York, 1962.
- [15] Nysten, B., Piraux, L., and Issi, J.-P., "Use of Thermal Conductivity Measurements as a Method to Characterize Carbon Fibers," , Pelnum Press, 1985, pp. 341-350.
- [16] Roth, E.P., Watson, R.D., Moss, M., and Drotning, W.D., "Thermophysical Properties of Advanced Carbon Materials for Tokamak Limiters," Sandia National Laboratories

SPIN DYNAMICS OF LAGEOS SATELLITE

**Arkady Kheyfets
Assistant Professor
Department of Mathematics**

**North Carolina State University
Raleigh, NC 27695-8205**

**Final Report for:
Summer Research Program
Phillips Laboratory**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.**

July 1992

SPIN DYNAMICS OF LAGEOS SATELLITE

Arkady Kheyfets
Assistant Professor
Department of Mathematics
North Carolina State University

Abstract

The long term prediction of the LAGEOS satellite (laser-ranged geodynamic satellite) spin dynamics is very important for a successful completion of its mission of measuring the relativistic Lense-Thirring effect. According to the current model, there is a possibility that, under combined influence of gravitational and magnetic fields of the Earth, the rotational dynamics of the LAGEOS satellite might become chaotic due to a fast growth of the satellite precession rate as its spin motion is slowed down by the Earth's magnetic field. We have reconsidered this prediction. A more detailed analysis of the Lagrange equation describing rotational dynamics of the satellite shows that the rate of the satellite precession is bounded, the bound being much smaller than the orbital angular velocity, even if magnetic forces are neglected. Inclusion of magnetic forces further reduces the value of the bound. Our Lagrangian formulation of the satellite spin dynamics can be used in the future investigation of the satellite nutation and the time scales of dissipating the rotational motions of the satellite under the influence of the gravitational and magnetic fields of the Earth.

SPIN DYNAMICS OF LAGEOS SATELLITE

Arkady Kheifets

INTRODUCTION

The LAGEOS satellite (laser-ranged geodynamic satellite) can be thought of, without getting into further details, as a sphere of radius ≈ 30 cm, outfitted with retroreflectors arranged in a way convenient for the laser ranging measurements. LAGEOS is in a nearly circular orbit 5900 km above the Earth's surface. Its orbit inclination is $\approx 110^\circ$, its orbital period is $\approx 1.35 \times 10^4$ s (≈ 3.75 hours)¹. Originally, LAGEOS was supposed to carry out the mission of measuring with high precision the geodynamic phenomena², such as polar motion, tides, plate motion, etc. Later, it was proposed to use LAGEOS together with another satellite LAGEOS-3 (to be launched) to measure the Lense - Thirring effect³ (also known under the name of inertial frame dragging), one of two effects predicted by general relativity but not yet observed. The Lense - Thirring effect predicts an additional to classical rotation of the orbital node of 31 mas/yr (milliarcsecond per year). To be able to measure this effect with desirable accuracy (under 10%) it is necessary to model geophysical effects on LAGEOS (this is supposed to be a part of the theory of the experiment) as well as satellite specific effects⁽⁴⁻⁶⁾.

The commonly accepted model of the satellite specific drag on the LAGEOS's orbital motion⁴ includes Yarkovsky thermal drag (70%), as well as the drag from neutral particles (14%) and charged particles (16%). Yarkovsky thermal drag depends strongly on the satellite spin axis position⁵ and can contribute ≈ 3 mas/yr to a rotation of the orbital node. Accordingly, a need in modeling of the spin dynamics of the satellite emerges. A recent model¹ takes into account the influence of the eddy currents induced in the aluminum core of LAGEOS by the Earth's magnetic field and the "Hipparcos" precession caused by the influence of gravitational gradient on the oblate spinning satellite. By now both effects caused by eddy currents (precession and dissipation of the spin rate) are confirmed by observation. On the contrary, the gravitational effect of "Hipparcos" precession potentially threatening to make the spin dynamics of LAGEOS chaotic has not been observed (at least not to the extent predicted by the theory)⁷. This discrepancy questions the ability of the current model¹ to make long term predictions of LAGEOS spin dynamics, and, as a consequence, of the satellite specific contribution to the nodal precession of the LAGEOS satellite. There is no doubt that

the discrepancy between the theoretical model and observational data is due to the assumptions of the model approximations .

Thus an additional investigation of the LAGEOS spin dynamics is necessary. The starting point of such an investigation should be a careful analysis of the current model assumptions and finding out the ones that are not adequate for describing the LAGEOS spin dynamics during the time period from present to the year 2000. This analysis should be followed by introducing appropriate corrections to the current model and/or a major overhaul of the model.

Our preliminary study indicates that there are several assumptions in the current model that are not likely to be satisfied under the described above conditions. We discover that even simple improvements of the model change drastically the long term predictions for the spin dynamics. The present report summarizes findings of our preliminary study of the problem and formulates a program for the future research in this important topic.

MAGNETIC TORQUE ON LAGEOS

As it has been mentioned above, the current model¹ of LAGEOS spin dynamics takes into account the influence of the eddy currents induced in the aluminum core of LAGEOS by the Earth's magnetic field and the precession caused by the influence of gravitational gradient on the oblate spinning satellite. The main acting torques in this model are due to (1) the eddy currents generated in the conducting body of the LAGEOS satellite by the $\vec{v} \times \vec{B}$ force due to the rotational velocity and (2) the gradient of the gravitational field of the Earth acting on the oblate distribution of LAGEOS's mass.

In this section we discuss the origin of the magnetic torques acting on LAGEOS.

In a rotating frame of the satellite the Earth's magnetic field is changing with the rotation frequency ω and induces in the body of the satellite eddy currents to a depth of the order of

$$\delta = \frac{c}{\sqrt{2\pi\omega\sigma}} \quad (1)$$

where c is the speed of light, and σ is the electrical conductivity ($\sigma = 2.2 \times 10^{17} \text{ s}^{-1}$)⁸. The induced current interacts with the Earth's magnetic field and produces the magnetic torque. In this report, when calculating magnetic effects, we assume that LAGEOS has the shape of a sphere. We neglect subtle details related to

the complex internal structure of the satellite because it does not seem to contribute to the drastic difference between our predictions and these of the current model. The magnetic torque exerted on the spherical conductor rotating with the angular velocity $\vec{\omega}$ is expressed by

$$\vec{\Gamma} = V\alpha'\hat{\omega} \times \vec{B}(\hat{\omega} \cdot \vec{B}) - V\alpha''(\vec{B} \times \hat{\omega}) \times \vec{B} = \vec{\Gamma}' + \vec{\Gamma}'' \quad (2)$$

where $\hat{\omega}$ is the unit vector pointing in the direction of $\vec{\omega}$, and $\alpha' + i\alpha'' = \alpha$ is the complex polarizability per unit volume (dimensionless), and V is the volume of the satellite. For the conducting sphere of radius ρ functions α' and α'' are given by⁹

$$\begin{aligned} \alpha' &= -\frac{3}{8\pi} \left[1 - \frac{3}{x} \frac{\sinh x - \sin x}{\cosh x - \cos x} \right] \\ \alpha'' &= -\frac{9}{4\pi} \frac{1}{x^2} \left[1 - \frac{1}{2} x \frac{\sinh x + \sin x}{\cosh x - \cos x} \right] \end{aligned} \quad (3)$$

where $x = \frac{2\rho}{c} = \frac{2\rho}{c} \sqrt{2\pi\omega\sigma}$. Two particular cases are of special importance, (1) the low frequency approximation ($x \ll 1$, or, equivalently, ω small) and the limit $\omega \rightarrow 0$ ($x \rightarrow 0$), and (2) the high frequency approximation ($x \gg 1$, or, equivalently, ω large) and the limit $\omega \rightarrow \infty$ ($x \rightarrow \infty$). For the low frequency approximation α' and α'' can be represented by series

$$\begin{aligned} \alpha' &= -\frac{1}{1680\pi} x^4 + \frac{1}{665280\pi} x^8 + \dots \\ \alpha'' &= -\frac{1}{80\pi} x^2 + \frac{1}{33600\pi} x^6 + \dots \end{aligned} \quad (4)$$

so that in the limit $\omega \rightarrow 0$ ($x \rightarrow 0$) both $\alpha' \rightarrow 0$ and $\alpha'' \rightarrow 0$. It is easy to figure out corresponding expressions for the high frequency approximation, particularly it is clear that, as $\omega \rightarrow \infty$ ($x \rightarrow \infty$), $\alpha' \rightarrow -\frac{3}{8\pi}$ and $\alpha'' \rightarrow 0$.

Bertotti and Iess¹ perform the magnetic torque analysis in the low frequency approximation. They argue that in this approximation the real part of polarizability α' is much smaller than its imaginary part α'' (cf. expressions (4)) and can be neglected, in which case the torque $\vec{\Gamma}$ can be approximated by

$$\vec{\Gamma} \approx \vec{\Gamma}'' = -V\alpha''\vec{\omega} \cdot (\vec{B}\vec{B} - \mathbf{I}B^2) \quad (5)$$

where \mathbf{I} is the unit dyadic. This torque obviously causes the decay of the satellite spin rate. More detailed analysis by Bertotti and Iess indicates that the direction of the spin also evolves and that the picture of the spin evolution can be rather involved.

MAGNETIC SLOWING DOWN OF LAGEOS

The magnetic field experienced by LAGEOS along its orbit is not constant. A reasonable suggestion¹ is to approximate the Earth's magnetic field by a dipole magnetic field

$$\vec{B} = \nabla \frac{\vec{d} \cdot \vec{r}}{r^3} = \frac{r^2 \vec{d} - 3\vec{r}(\vec{r} \cdot \vec{d})}{r^5} \quad (6)$$

where \vec{d} is the Earth's magnetic dipole vector. Since the angular velocity changes very little in an orbital period it is reasonable to average the torque (5) over the orbit. In the frame where the node is at rest and the components of the unit vector parallel to the Earth's axis are (0, 0, 1) while the components of the unit vector orthogonal to the orbit plane are (sin I , 0, cos I) (where I is the orbital inclination) the averaging procedure leads to the positive definite matrix

$$\langle B^2 \delta_{ij} - B_i B_j \rangle = \frac{d^2}{a^6} \beta_{ij} \quad (7)$$

where a is the orbital radius, and β_{ij} is the matrix

$$\beta_{ij} = \begin{pmatrix} b_1 & 0 & b_2 \\ 0 & b_3 & 0 \\ b_2 & 0 & b_4 \end{pmatrix} \quad (8)$$

with

$$\begin{aligned} b_1 &= \frac{1}{8}(20 - 39 \cos^2 I + 27 \cos^4 I) \\ b_2 &= \frac{3}{8}(5 - 9 \cos^2 I) \cos I \sin I \\ b_3 &= \frac{1}{8}(11 - 3 \cos^2 I) \\ b_4 &= \frac{3}{8}(3 - 4 \cos^2 I - 9 \cos^4 I) \end{aligned} \quad (9)$$

The averaged magnetic torque causes the mean rotation vector ω_i to evolve according to the equation

$$\frac{1}{\nu_m} \frac{d\omega_i}{dt} = -\beta_{ij} \omega_j \quad (10)$$

where

$$\nu_m = V \frac{\alpha''}{I_3} \frac{a^2}{d^6} \quad (11)$$

is the value of the magnetic spin decay frequency (I_3 is the moment of inertia with respect to the principal axis parallel to $\vec{\omega}$). The ν_m is the measure of how fast the LAGEOS's spin decays. One has to keep in mind that, in such a description, the actual evolution of spin (including its direction) can be rather involved

and depends crucially on the eigenvalues and eigenvectors of the matrix β_{ij} . One can find a rather detailed description of the different possibilities in the paper by Bertotti and Iess¹. We are not considering these details here. Although it is an interesting subject, it has nothing to do, in our opinion, with the long term prediction of LAGEOS's spin dynamics. Any attempt to extrapolate this technique to the case of small spin will fail. The breakdown will take place long before the averaging procedure will lose its meaning. As we will see the basic procedures can be applied for the small value of spin. The main problem with the described technique is an implicit assumption of a rigid split of rotation into the spin itself and precession that cannot be sustained at small values of the spin rate. However this assumption can be dropped without essential losses if the goal is restricted to asymptotic predictions.

GRAVITATIONAL PRECESSION: FAST SPINNING SATELLITE

Gravitational precession is caused by the influence of gravitational gradient on the oblate spinning satellite. The oblateness of the satellite is determined by the combination

$$\Delta = \frac{I_3 - I_1}{I_3} \quad (12)$$

where I_3 and $I_1 = I_2$ are the principal moments of inertia (the principal direction corresponding to I_3 is assumed to be that of $\vec{\omega}$ (coinciding with the direction of the symmetry axis of the oblate satellite)). The gradient of the Earth's gravitational field acts on the oblate mass distribution of LAGEOS and produces a precession of its spin vector around the normal to the orbital plane at a rate ω_p given by the formula

$$\omega_p = \frac{3}{2} \Delta \frac{\omega_0^2}{\omega_3} \cos \theta \quad (13)$$

for $\omega_p \ll \omega_0 \ll \omega$, where ω_0 is the orbital angular velocity, ω_3 is the satellite spin, and θ is the obliquity angle of spacecraft (the angle between $\vec{\omega}$ and the normal to the orbital plane). This gravitational precession has the same mechanism as the lunisolar Hipparcos precession of the Earth¹⁰. The value of the LAGEOS's oblateness is $\Delta \approx 3.35\%$ ¹. The initial gravitational precession period was of the order of hundreds of years but should decrease rapidly as spin motion of the satellite slows down.

It is argued by Bertotti and Iess¹ that, since $\omega_p \propto 1/\omega$ the gravitational precession in the asymptotic limit dominates and may make the spin dynamics chaotic. According to their original prediction the first signs of such a behavior should show up by now. This prediction has not been confirmed by observations

and, as we will argue, is not likely to be observed in far future. It is based on the equation (13) which can be applied only if the spin rate $|\omega_3|$ is large. It is assumed that eventually the period of gravitational precession will become very short and it is concluded¹ that LAGEOS, like a very slow top, will start tumbling faster and faster, with a chaotic dynamics until eventually the scale of variations of $\vec{\omega}$ will become shorter than the orbital period of the satellite and the original model of Bertotti and Iess will be no longer applicable.

We believe that the latter prediction of Bertotti and Iess is too pessimistic. Their model, with minor improvements, is much more viable than it might appear. It is not quite clear what is the nature of the predicted tumbling. Perhaps, it is nutation. If so, it can be included in the model without too much trouble and analyzed by relatively simple means. As about breaking the averaging approximation because of fast precession and nutation, the situation is also not so desperate. As we will see later it simply does not happen.

APPROXIMATIONS IN LAGEOS SPIN DYNAMICS

A discrepancy between the theoretical predictions concerning gravitational precession¹ and the recent observational evidence presents us with a strong motivation to reconsider carefully approximations made in the process of deriving basic equations describing LAGEOS's spin dynamics and the adequacy of these approximations in the case of LAGEOS. These approximations are necessary not only to make the problem tractable but also, and this is much more important, to make sure that a model captures sufficient features of the phenomenon under consideration and ignores the unimportant aspects of the problem, the ones that seem to be important to provide a better fit with available experimental data but in fact are unnecessary and eventually prevent the model from making correct predictions on the future evolution of the system. The issue of correct approximations is especially important in problems demanding long term predictions, like the LAGEOS's spin dynamics.

The assumptions in modeling the LAGEOS's spin dynamics can be subdivided in several different classes. These are:

(1) Assumptions on the nature of forces acting on the satellite.

These are the assumptions that enable one to write down the expressions for potential functions, Lagrangians, torques, etc. Ordinarily, they include some assumptions concerning the properties of the materials, shapes of the objects, influence of motions of the objects on these properties. Any

change in these assumptions usually implies a very extensive effort. We are going to avoid doing it for as long as it does not appear absolutely necessary.

(2) Assumption concerning the independence of degrees of freedom.

This assumption is present implicitly in the Bertotti-less model and expresses itself in neglecting the nutations. It prevents the model from analyzing the transition to the regime of tumbling. We will discard this assumption eventually. This requires introduction of the full system of Lagrange equations and leads to some complications but the problem appears to be tractable for the asymptotic analysis. Neglecting the nutations can also lead to incorrect conclusions concerning precessions because nutations and precessions are coupled.

(3) Assumptions on the relative rates of the orbital angular velocity, spin, and precession.

The model of Bertotti and less uses the standard assumption that the rate of precession is much smaller than the orbital angular velocity. This assumption allows the to use the averaging procedure (averaging over the orbit) for the magnetic torque and, implicitly, for the Earth's gravitational field gradient. This assumption appears to lead to breakdown of the model as soon as the satellite spin slows down sufficiently (cf. equation (13)) and the rate of precession exceeds the orbital angular velocity, which would make the averaging procedure meaningless. It is interesting that the potentially troublesome averaging procedure turns out to be rather harmless. The reason for this is that formula (13) is correct only for large values of spin. At small values of spin the rate of precession is given by a different expression (cf. next section) according to which the rate of precession is bounded by the value much smaller than the orbital angular velocity. The assumption in question, and together with it the orbital averaging procedure, can be, unexpectedly, retained with no severe penalty for as long as the rate of precession is being calculated (but should be dropped when discussing nutations). The assumption of the spin being large compared to the orbital angular velocity is completely different matter. It is a very common assumption in the rigid body mechanics. It is used in the theory of the equinoxes precession, and, in particular, in deriving expression (13). It is absolutely unnecessary, cannot be satisfied in the study of asymptotic behavior of LAGEOS's spin, and dropping it only improves the physics of the model. We will illustrate this point in the next section.

In conclusion of this section a few general remarks are in order. It is very hard to believe predictions of the current model¹ concerning asymptotic behavior of the LAGEOS's spin dynamics at $t \rightarrow \infty$ ($\omega_3 \rightarrow 0$).

- (i) For any mechanism of spin dissipation it seems hardly possible that the precession frequency can grow too much. Energy associated with the precession motion is coming from the potential energy of the oblate mass distribution of the satellite in the inhomogeneous gravitational field of Earth and is bounded which implies that the precession rate should be bounded too.
- (ii) The same mechanism as that of spin dissipation (eddy currents) should also dissipate motions associated with precession and nutation. Admittedly, in the beginning the rate of precession is very small compared with the spin. However this cannot be told about nutation. A rough estimate of the nutation frequency (using the formulae derived for the fast top) at large values of spin leads to

$$a = \frac{I_3}{I_1} \omega_3 \approx \omega_3 \quad (14)$$

which means that the nutation frequency is of the same order of magnitude as the spin rate and, as a consequence, the rate of decay of nutation should be about the same as that of spin. Under these conditions one cannot expect the satellite to tumble eventually like a very slow top. As about the decay of precessions, it should eventually occur, too, although the time scale of this decay might be different.

More generally, it should be, probably, expected that all spin related motions will be eventually decayed except those that are sustained by some constantly present torque. Roughly speaking, the situation should be like in the case of Moon orbiting around Earth, although the decay mechanism in the latter case is quite different from this of LAGEOS.

GRAVITATIONAL PRECESSION: ARBITRARY SPIN OF THE SATELLITE

In this section we are going to analyze the procedure of calculating the gravitational precession of LAGEOS dropping the assumption of the fast rate of spin, but retaining the assumption of the rate of precession being slow compared to the orbital angular velocity. This analysis is much simpler than the analysis of the full system of Lagrange equations describing the spin dynamics of LAGEOS interacting with both magnetic and gravitational field of the Earth, it gives an idea on the origin of the bounds imposed on the parameters of the spinning satellite, and motivate the further study of the complete system of equations. In the next section we describe briefly the most important results of our study of asymptotics of the LAGEOS's spin dynamics.

The equations governing the LAGEOS's gravitational precession are the same as those describing the Earth's precession about the normal to the ecliptic, a motion known in astronomy as the precession of equinoxes, or Hipparchos precession¹⁰. In our case this precession is caused by the action of the Earth's gravitational field on the slightly oblate LAGEOS. To calculate the rate of this precession one needs to calculate the mutual gravitational potential of the Earth, approximated in the standard procedure of derivation¹⁰ by a mass point M , and a nonspherical distribution of matter representing LAGEOS. Approximating the Earth by a mass point is equivalent to neglecting the Earth's oblateness which appears quite adequate in matters concerning the satellite spin dynamics, as simple considerations indicate that the Earth's oblateness coupled only to orbital parameters in the lowest order approximation.

Evaluation of potential energy is performed via expansion of energy in a power series with coefficients proportional to appropriate Legendre polynomials and subsequent truncation neglecting the terms of higher than the second power. It is assumed in calculations that the axis of symmetry is the third principal axis of the LAGEOS's inertia tensor, so that $I_1 = I_2 \neq I_3$. If α, β, γ are the direction cosines of the radius vector from the LAGEOS's center to the Earth with respect to the principal axes, then the expression for the potential is given by

$$V = -\frac{GMm}{r} + \frac{GM(I_3 - I_1)}{2r^3} [3\gamma^2 - 1], \quad (15)$$

or

$$V = -\frac{GMm}{r} + \frac{GM(I_3 - I_1)}{r^3} P_2(\gamma). \quad (16)$$

where $P_2(x)$ is the second Legendre polynomial determined by

$$P_2(x) = \frac{1}{2}(3x^2 - 1). \quad (17)$$

In what follows we assume LAGEOS to be rigid. In particular, we require that the principal axes be fixed in LAGEOS and the associated moments of inertia be constant in time.

In expression (16) for the potential only the second term

$$V_2 = \frac{GM(I_3 - I_1)}{r^3} P_2(\gamma) \quad (18)$$

depends on the orientation of LAGEOS and thus could give rise to torques. It was mentioned above that γ is the direction cosine between the LAGEOS axis and the radius vector from the LAGEOS's center to the Earth. As the Earth goes around its apparent orbit γ is changing. It is customary to express γ in terms of

functions of two angles, the first one (θ) being the angle between the axis of LAGEOS and the normal to the LAGEOS's orbit plain (also called the obliquity of the LAGEOS's axis), the second one (η) being the angular position of the Earth on its apparent orbit. The relation between γ , θ , and η is

$$\gamma = \sin \theta \cos \eta. \quad (19)$$

Hence, V_2 can be written

$$V_2 = \frac{GM(I_3 - I_1)}{2r^3} [3 \sin^2 \theta \cos^2 \eta - 1]. \quad (20)$$

In typical applications the orbital motion is very rapid compared to the precessional motion. The usual practice is to average V_2 over a complete orbital period for the purpose of obtaining the mean precession rate. The orbit of LAGEOS has very low eccentricity, which implies that r can be assumed constant and the only variation is in $\cos \eta$ (with $\overline{\cos^2 \eta} = \frac{1}{2}$), so that

$$\overline{V_2} = \frac{GM(I_3 - I_1)}{2r^3} \left[\frac{3}{2} \sin^2 \theta - 1 \right] = \frac{GM(I_3 - I_1)}{2r^3} \left[\frac{1}{2} - \frac{3}{2} \cos^2 \theta \right]. \quad (21)$$

Finally, the expression for $\overline{V_2}$ takes form

$$\overline{V_2} = \frac{GM(I_3 - I_1)}{2r^3} P_2(\cos \theta). \quad (22)$$

The torque derived from equation (22) is perpendicular to both LAGEOS's axis and the normal to the orbit (which plays the same role as the vertical axis for the heavy top). Hence the precession is about the direction of the orbit normal vector. The magnitude of the precession rate can be obtained in several different ways. We choose to use Lagrange equations approach because it makes assumptions of an accepted approximation transparent.

For a symmetric body in which the potential is a function of $\cos \theta$ only, as it is the case with LAGEOS, the Lagrangian can be written as

$$L = \frac{I_1}{2} (\dot{\theta}^2 + \dot{\phi}^2 \sin^2 \theta) + \frac{I_3}{2} (\dot{\psi} + \dot{\phi} \cos \theta)^2 - V(\cos \theta). \quad (23)$$

At this point we are not going to write the full Lagrangian equation corresponding to θ . Instead, as it is done ordinarily, we assume only uniform precession and do not concern ourselves about the necessary initial conditions. We simply take $\dot{\theta} = 0$ and $\ddot{\theta} = 0$ in the equations of motion. The simplified Lagrangian equation corresponding to θ is then

$$\frac{\partial L}{\partial \theta} = I_1 \dot{\phi}^2 \sin \theta \cos \theta - I_3 \dot{\phi} \sin \theta (\dot{\psi} + \dot{\phi} \cos \theta) - \frac{\partial V}{\partial \theta} = 0. \quad (24)$$

Taking into account that

$$\omega_3 = \dot{\psi} + \dot{\phi} \cos \theta, \quad (25)$$

and, for $V = V(\cos \theta)$,

$$\frac{\partial V}{\partial(\cos \theta)} = -\frac{1}{\sin \theta} \frac{\partial V}{\partial \theta}, \quad (26)$$

we can rewrite equation (24) in the form

$$I_3 \omega_3 \dot{\phi} - I_1 \dot{\phi}^2 \cos \theta = \frac{\partial V}{\partial(\cos \theta)}. \quad (27)$$

In our case

$$\frac{\partial V}{\partial(\cos \theta)} = \frac{\partial \bar{V}_2}{\partial(\cos \theta)} = -\frac{3}{2} \frac{GM(I_3 - I_1)}{r^3} \cos \theta. \quad (28)$$

Using Kepler's law

$$\omega_0^2 = \left(\frac{2\pi}{\tau}\right)^2 = \frac{GM}{r^3}, \quad (29)$$

where ω_0 the angular velocity of the orbital motion and τ is the orbital period, we can write instead of (28)

$$\frac{\partial V}{\partial(\cos \theta)} = \frac{\partial \bar{V}_2}{\partial(\cos \theta)} = -\frac{3}{2} \omega_0^2 (I_3 - I_1) \cos \theta. \quad (30)$$

Hence, the Lagrange equation (27) takes the following final form

$$I_3 \omega_3 \dot{\phi} - I_1 \dot{\phi}^2 \cos \theta = -\frac{3}{2} \omega_0^2 (I_3 - I_1) \cos \theta. \quad (31)$$

Standard applications of equation (31), such as, for instance, the precession of the Earth's equinoxes, assume also that the precession rate $\dot{\phi}$ is small. It is usually stated¹⁰ that for the case $\dot{\phi} \ll \omega_3$ the $\dot{\phi}^2$ term in equation (31) can be dropped. As we will see later this procedure (dropping the term quadratic in $\dot{\phi}$) is not as innocent as it might look from the first sight. But for the time being we will go along with the standard line of reasoning. Neglecting this quadratic term immediately leads to the following expression for the rate of precession

$$\dot{\phi} = -\frac{3}{2} \frac{\omega_0^2}{\omega_3} \frac{I_3 - I_1}{I_3} \cos \theta = -\frac{3}{2} \frac{\omega_0^2}{\omega_3} \Delta \cos \theta, \quad (32)$$

where $\Delta = (I_3 - I_1)/I_3$ determines the oblateness of LAGEOS. This expression is used to conclude¹ that the rate of gravitational precession is proportional to the period of spin motion of the satellite $T = \frac{2\pi}{\omega_3}$. Furthermore, it is stated that when the period T increases exponentially under the influence of the Earth's

magnetic field, so does the rate of gravitational precession, and that this rate of precession can become very large threatening to lead the spin dynamics of LAGEOS to chaos ($|\dot{\phi}| \rightarrow \infty$ when $\omega_3 \rightarrow 0$).

A rather simple analysis shows that this conclusion is wrong. A misunderstanding emerges entirely due to an attempt to use equation (32) in the domain where it is not applicable. The expression is appropriate only for small magnitudes of $\dot{\phi}$ (which is indeed the case in describing, say, the precession of the Earth's equinoxes, but not in analyzing the LAGEOS's spin dynamics). More precisely, it can be used only when the term quadratic in $\dot{\phi}$ can be dropped from the Lagrange equation (31). To see what is involved in such an assumption, we return to equation (31) and rewrite it in the following form

$$\dot{\phi}^2 - \frac{I_3 \omega_3}{I_1 \cos \theta} \dot{\phi} - \frac{3}{2} \omega_0^2 \frac{I_3 - I_1}{I_1} = 0, \quad (33)$$

or

$$\dot{\phi}^2 - C_1 \dot{\phi} - C_2 = 0, \quad (34)$$

where

$$\begin{aligned} C_1 &= \frac{I_3}{I_1} \frac{\omega_3}{\cos \theta} \approx \frac{\omega_3}{\cos \theta} \\ C_2 &= \frac{3}{2} \omega_0^2 \frac{I_3 - I_1}{I_1} \approx \frac{3}{2} \omega_0^2 \Delta \end{aligned} \quad (35)$$

The solution of the quadratic equation (34) is

$$\dot{\phi} = \frac{C_1 \pm \sqrt{C_1^2 + 4C_2}}{2} = \frac{1}{2} \left(\frac{\omega_3}{\cos \theta} \pm \sqrt{\frac{\omega_3^2}{\cos^2 \theta} + 6\omega_0^2 \Delta} \right). \quad (36)$$

A simple reasoning (no oblateness, no precession) resolves the ambiguity of the sign in front of the square root (the minus sign should be taken), which leaves us with the solution

$$\dot{\phi} = \frac{1}{2} \left(\frac{\omega_3}{\cos \theta} - \sqrt{\frac{\omega_3^2}{\cos^2 \theta} + 6\omega_0^2 \Delta} \right). \quad (37)$$

This solution has finite limits in both cases when $\omega_3 \rightarrow \infty$ and when $\omega_3 \rightarrow 0$

$$\begin{aligned} \lim_{\omega_3 \rightarrow \infty} \dot{\phi} &= 0, & \lim_{\omega_3 \rightarrow 0} \dot{\phi} &= -\omega_0 \sqrt{\frac{3}{2} \Delta}, \\ \Delta \ll 1 &\implies \left| \lim_{\omega_3 \rightarrow 0} \dot{\phi} \right| \ll |\omega_3|. \end{aligned} \quad (38)$$

and at any value of $0 < \frac{1}{\omega_3} < \infty$ the value of the solution belongs to the interval determined by (38). When T increases exponentially the rate of precession $\dot{\phi}$ does not follow it.

It is instructive to rewrite equation (37) in form

$$\dot{\phi} = \frac{1}{2} \frac{\omega_3}{\cos \theta} \left(1 - \sqrt{1 + \frac{6\omega_0^2 \Delta \cos^2 \theta}{\omega_3^2}} \right), \quad (39)$$

and look at its approximation when $\frac{6\omega_0^2 \Delta \cos^2 \theta}{\omega_3^2}$ is small. It is easy to see that in this case

$$\dot{\phi} \approx \frac{1}{2} \frac{\omega_3}{\cos \theta} \left(1 - 1 - \frac{1}{2} \frac{6\omega_0^2 \Delta \cos^2 \theta}{\omega_3^2} \right) \approx -\frac{3}{2} \frac{\omega_0^2}{\omega_3} \Delta \cos \theta, \quad (40)$$

and we recover expression (32). However, the expansion of the square root that we use to recover this expression explicitly demands smallness of the ratio

$$\frac{6\omega_0^2 \Delta \cos^2 \theta}{\omega_3^2}. \quad (41)$$

A careful look at (41) uncovers that the procedure of dropping the term proportional to $\dot{\phi}^2$ in equation (31) not only demands smallness of $\dot{\phi}$ but also, at fixed values of θ , Δ , and ω_0 , imposes the lower limit on the magnitude of ω_3 . In other words, in order for (32) to work the satellite should have enough spin. Equations (37), (38) also indicate that at small values of ω_3 neglecting ω_1 and ω_2 would not be appropriate. However, an assumption of smallness of ω_1 and ω_2 still can be used in neglecting their higher powers, provided that it is done carefully.

LAGRANGE EQUATIONS FOR THE LAGEOS SPIN DYNAMICS

The results of the previous section indicate that the LAGEOS's spin dynamics at slow spin rate might look quite different from what one expects making intuitive assumptions coming from experience in high spin frequency domain. A safer approach would be to obtain first the full system of equations governing the spin dynamics and to introduce simplifying assumptions on the later stage, depending on the question asked. As in previous section, we are going to use the Lagrange formulation of the problem. Restrictions on the length of this report prevent us from providing most of the details of the calculations. We are going to review only the basic principles and some of the preliminary conclusions. Some additional details can be found in the report of C. Fuchs¹¹ in this volume.

As in previous section, we introduce the frame of principal axes of inertia (the body-fixed frame) rotation of which with respect to space-fixed frame is described by the Euler angles θ , ϕ , and ψ . The

components of the angular velocity $\vec{\omega}$ in the body-fixed frame are

$$\begin{aligned}\omega_1 &= \dot{\phi} \sin \theta \sin \psi + \dot{\theta} \cos \psi \\ \omega_2 &= \dot{\phi} \sin \theta \cos \psi - \dot{\theta} \sin \psi \\ \omega_3 &= \dot{\phi} \cos \theta + \dot{\psi}\end{aligned}\quad (42)$$

To include in the dynamic picture the action of the Earth's magnetic field on the satellite one needs to add to the gravitational Lagrangian

$$L_g = \frac{I_1}{2}(\dot{\theta}^2 + \dot{\phi}^2 \sin^2 \theta) + \frac{I_3}{2}(\dot{\psi} + \dot{\phi} \cos \theta)^2 - V \quad (43)$$

the term $L_m = \vec{M} \cdot \vec{B}$, where \vec{B} is the external magnetic field and \vec{M} is the magnetic moment induced by \vec{B} in the satellite (V in equation (43) represents gravitational potential energy determined by either expression (20) or (22) depending on whether one tries to evaluate instantaneous quantities or the ones averaged over the orbital period). In other words the full Lagrangian for the LAGEOS's spin dynamics is

$$L = L_g + L_m = \frac{I_1}{2}(\dot{\theta}^2 + \dot{\phi}^2 \sin^2 \theta) + \frac{I_3}{2}(\dot{\psi} + \dot{\phi} \cos \theta)^2 - V + \vec{M} \cdot \vec{B}. \quad (44)$$

To calculate the L_m one needs to use the relation between the components of magnetic field B_1, B_2, B_3 in the body-fixed frame and $\bar{B}_1, \bar{B}_2, \bar{B}_3$ in the space-fixed frame

$$\begin{aligned}B_1 &= (\cos \psi \cos \phi - \cos \theta \sin \phi \sin \psi) \bar{B}_1 + (\cos \psi \sin \phi + \cos \theta \cos \phi \sin \psi) \bar{B}_2 + (\sin \psi \sin \theta) \bar{B}_3 \\ B_2 &= (-\sin \psi \cos \phi - \cos \theta \sin \phi \cos \psi) \bar{B}_1 + (-\sin \psi \sin \phi + \cos \theta \cos \phi \cos \psi) \bar{B}_2 + (\cos \psi \sin \theta) \bar{B}_3 \\ B_3 &= (\sin \theta \sin \phi) \bar{B}_1 - (\sin \theta \cos \phi) \bar{B}_2 + (\cos \theta) \bar{B}_3\end{aligned}\quad (45)$$

and the relation between the magnetic field \vec{B} , polarizability⁹ α (cf. equation (3)), and the induced magnetic moment \vec{M} . Equation (3) indicates that α is a function of the frequency ω . The equations below will depend on three frequencies $\omega_1, \omega_2, \omega_3$. The notation α_k will refer to $\alpha(\omega_k)$ for $k = 1, 2, 3$. The resulting expression for L_m is

$$L_m = \vec{M} \cdot \vec{B} = V [\alpha'_1(B^2 - B_1^2) + \alpha'_2(B^2 - B_2^2) + \alpha'_3(B^2 - B_3^2)], \quad (46)$$

where $B^2 = B_1^2 + B_2^2 + B_3^2 = \bar{B}_1^2 + \bar{B}_2^2 + \bar{B}_3^2$.

One could write now the full system of Lagrange equations of the satellite spin dynamics

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} = 0, \quad \frac{d}{dt} \frac{\partial L}{\partial \dot{\phi}} - \frac{\partial L}{\partial \phi} = 0, \quad \frac{d}{dt} \frac{\partial L}{\partial \dot{\psi}} - \frac{\partial L}{\partial \psi} = 0. \quad (47)$$

An investigation of such equations (both analytically and numerically) is under way at present time but can be completed only in course of the follow-up research. In any case, the full discussion of the results (even those available now) would be beyond the scope of this report. Instead, we are going to describe a qualitative result similar to the one of the previous section.

We are going to obtain expressions for precession, averaged over the orbit, extending the results of previous section to the combined effect of both magnetic and gravitational forces. Exactly as in previous section, we assume $\dot{\theta} = 0$ and $\ddot{\theta} = 0$ (this assumption will be discussed later). The resulting Lagrange equation for θ takes form (compare to equation (31))

$$I_1 \dot{\phi}^2 \sin \theta \cos \theta - I_3 \omega_3 \dot{\phi} \sin \theta - \frac{3}{2} \omega_0^2 (I_3 - I_1) \sin \theta \cos \theta + \frac{\partial}{\partial \theta} (\vec{M} \cdot \vec{B}) - \frac{d}{dt} \frac{\partial}{\partial \dot{\theta}} (\vec{M} \cdot \vec{B}) = 0 \quad (48)$$

The last two terms of this equation describe coupling of spin degrees of freedom to the Earth's magnetic field. For their evaluation or estimate it is necessary to be able to calculate the expressions (cf. previous sections for notations)

$$\begin{aligned} \frac{\partial \alpha'_k}{\partial \theta} &= \frac{\partial \alpha'_k}{\partial x_k} \frac{\partial x_k}{\partial \omega_k} \frac{\partial \omega_k}{\partial \theta} = \frac{\rho}{c} \sqrt{2\pi\sigma} \frac{1}{\omega_k} \frac{\partial \alpha'_k}{\partial x_k} \frac{\partial \omega_k}{\partial \theta} = f(x_k) \frac{\partial \omega_k}{\partial \theta} \\ \frac{\partial \alpha'_k}{\partial \theta} &= \frac{\partial \alpha'_k}{\partial x_k} \frac{\partial x_k}{\partial \omega_k} \frac{\partial \omega_k}{\partial \theta} = \frac{\rho}{c} \sqrt{2\pi\sigma} \frac{1}{\omega_k} \frac{\partial \alpha'_k}{\partial x_k} \frac{\partial \omega_k}{\partial \theta} = f(x_k) \frac{\partial \omega_k}{\partial \theta} \end{aligned} \quad (49)$$

where

$$f(x) = \frac{4\pi\rho^4}{c^2} \sigma \frac{1}{x} \frac{\partial \alpha'}{\partial x} = \frac{9\rho^2\sigma}{c^2} \frac{1}{x^2} \left[2 \frac{1 - \cosh x \cos x}{(\cosh x - \cos x)^2} - \frac{1}{x} \frac{\sinh x - \sin x}{\cosh x - \cos x} \right] \quad (50)$$

Since we are planning to investigate the domain of small ω it is useful to stress that, for small ω

$$\alpha' \sim \omega^2, \quad \frac{d\alpha'}{dx} \sim \omega^{\frac{3}{2}}, \quad \frac{d\alpha'}{d\omega} \sim \omega, \quad f(x) \sim \omega \quad (51)$$

Computations for $\frac{\partial}{\partial \theta} (\vec{M} \cdot \vec{B})$ and $\frac{\partial}{\partial \theta} (\vec{M} \cdot \vec{B})$ yield

$$\begin{aligned} \frac{\partial}{\partial \theta} (\vec{M} \cdot \vec{B}) &= V [(B^2 - B_1^2) f(x_1) \dot{\phi} \cos \theta \sin \psi + (B^2 - B_2^2) f(x_2) \dot{\phi} \cos \theta \cos \psi \\ &\quad - (B^2 - B_3^2) f(x_3) \dot{\phi} \sin \theta - 2(\alpha'_1 \frac{\partial B_1}{\partial \theta} + \alpha'_2 \frac{\partial B_2}{\partial \theta} + \alpha'_3 \frac{\partial B_3}{\partial \theta})] \end{aligned} \quad (52)$$

and

$$\begin{aligned} \frac{\partial}{\partial \theta} (\vec{M} \cdot \vec{B}) &= V \left[(B^2 - B_1^2) \frac{\partial \alpha'_1}{\partial \theta} + (B^2 - B_2^2) \frac{\partial \alpha'_2}{\partial \theta} + (B^2 - B_3^2) \frac{\partial \alpha'_3}{\partial \theta} \right] \\ &= V [(B^2 - B_1^2) f(x_1) \cos \psi - (B^2 - B_2^2) f(x_2) \sin \psi] \end{aligned} \quad (53)$$

At this point we undertake a shortcut. Keeping in mind that we are trying to compute the rate of the satellite precession when the satellite rotates slowly and spins around its axis of symmetry even slower we

drop in the Lagrange equation (48) the terms of the first order with respect to ω_3 (and higher orders), and the terms of the second order (and higher) with respect to ω_2, ω_3 . Furthermore, as we are trying to compute only the (quasi)stationary part of precession, we drop all the terms proportional to the angular accelerations $\dot{\omega}_k$. The final form of the Lagrange equation under these assumptions is

$$\dot{\phi}^2 - C \dot{\phi} - \frac{3}{2} \omega_0^2 \Delta = 0 \quad (54)$$

where

$$C = -\frac{V}{I_1 \sin \theta} [(B^2 - B_1^2)f(x_1) \sin \psi + (B^2 - B_2^2)f(x_2) \cos \psi] \quad (55)$$

so that the mean value of the satellite precession rate is given by

$$\dot{\phi} = \frac{1}{2}C - \sqrt{\frac{C^2}{4} + \frac{3}{2}\omega_0^2\Delta} \quad (56)$$

The solution for $\dot{\psi}$ has some attractive features: (i) it has the correct limit coinciding with (38) when the magnetic field is switched off, and (ii) it predicts that the mean precession rate is slower in presence of magnetic field than it would be without it. However, the solution is not quite stationary (which shows up in the dependence of C on ψ). This indicates that the demand to have a stationary solution might be not quite compatible with dynamics determined by the total Lagrangian (44) of the system, or that a more sophisticated averaging procedure for the magnetic part of the Lagrangian is needed. In any case, exactly like in previous section, all indications are that the precession rate of the satellite is bounded by the limit well under the orbital angular velocity of the satellite. It is hardly probable that the chaotic spin dynamics can emerge through this mechanism. Another possibility of transition to the chaotic dynamical regime, namely through the growth of the satellite nutation, requires more sophisticated treatment and should be investigated in the future.

CONCLUSIONS

In this report we have described the results of our research on the LAGEOS's spin dynamics. We have reconsidered some conclusions of the current model, particularly its prediction concerning a transition of the rotational motion of the satellite to a chaotic regime due to the rapid increase of the precession rate as the spinning motion around the satellite symmetry axis slows down. We have analyzed assumptions underlying this conclusion of the current model and have established that some of these assumptions are not

appropriate for describing the satellite spin dynamics at low spin rates. We have dropped these assumptions and came to the conclusion that even in absence of magnetic fields the rate of precession is bounded, the bound being much smaller than the orbital angular velocity of the satellite for the low value of the satellite oblateness. This conclusion has motivated our further work. We have developed the Lagrangian formulation of the satellite spin dynamics, included in it the term describing the coupling of the satellite to the Earth's magnetic field via eddy currents mechanism. Using the obtained Lagrange equations, we have analyzed qualitatively the mean precession rate of the satellite coupled to both gravitational and magnetic field of the Earth. The resulting expression for the satellite precession rate agrees with the expression for the precession rate bound in pure gravitational field when magnetic coupling is switched off and predicts the value of the bound smaller than that for pure gravitational coupling. We have come to the conclusion that the technique of averaging over the orbital period can be sustained in a broader domain of conditions than it was believed, for as long as precessions are the subject of interest.

Another possible way for the satellite to transit to the chaotic regime of rotational motion is through the growth of the nutation amplitude when the satellite spinning motion slows down, similarly to what happens to the spinning heavy top. However this analogy is, in our opinion, dangerous. The "magnetic friction" mechanism created by eddy currents slows down any rotational motion at a rate depending on the frequency of this motion, and the nutation frequency is of the same order of magnitude as the spin frequency, at least right after the satellite launch. It should be acted upon by the "magnetic friction" strongly, and it is not clear whether it will ever lead to "tumbling" of the satellite. Additional investigation of this problem is required. Most probably, it will require an effort combining both analytic and numerical techniques.

In addition to investigation of the asymptotic behavior of the Lagrange equations solution, an important question to consider is the time scales of dissipation for different rotational degrees of freedom. In other words, the question is how close is the satellite spin dynamics to asymptotic within the period of measurements related to the Lense-Thirring effect. Again, a combination of analytic and numerical efforts will be required to answer this important question.

It is quite possible that the orbital averaging technique will fail when we come to all these questions. In this case numerical methods will have to be relied upon more heavily in the future investigations. But, whatever the means are, it is clear that the efforts in studying the LAGEOS's spin dynamics should continue until all the question are answered.

REFERENCES

- ¹ Bertotti B., Iess L., The rotation of LAGEOS, *J. Geophys. Research*, **96**, 2431 (1991).
- ² *J. Geophys. Research*, (the LAGEOS special section) **90**, 9215 (1985).
- ³ Ciufolini I., Measurement of the Lense-Thirring drag effect on LAGEOS and another high altitude laser ranged satellite, *Phys. Rev. Lett.*, **56**, 278 (1986).
- ⁴ Rubincam D. P., Drag on the LAGEOS satellite, *J. Geophys. Research*, **95**, 4881 (1990).
- ⁵ Rubincam D. P., Yarkovski thermal drag on LAGEOS, *J. Geophys. Research*, **93**, 13805 (1988).
- ⁶ Rubincam D. P., and Weiss N. R., The orbit of LAGEOS and solar eclipses, *J. Geophys. Research*, **90**, 9399 (1985).
- ⁷ Miller W. A., Private communication (1992).
- ⁸ Harper C. A., *Handbook of material and processes for electronics*, MacGraw-Hill, NY (1970).
- ⁹ Landau L.D., Lifshitz E. M., *Electrodynamics of continuous media*, Pergamon Press, Oxford (1975).
- ¹⁰ Goldstein H., *Classical mechanics*, Addison-Wesley Publ. Co., Reading, MA (1981).
- ¹¹ Fuchs C., *Lagrangian formulation of LAGEOS's spin dynamics*, p. 7-1 of this volume

**Generation of ELF and VLF Waves by a Thermal Instability Excited in the HF
Heater-Modulated Polar Electrojet**

S. P. Kuo
Professor
Department of Electrical Engineering

Polytechnic University
Route 110
Farmingdale, NY 11735

Final Report for:
Summer Research Project
Geophysics Directorate of the Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

August, 1992

Generation of ELF and VLF Waves by a Thermal Instability Excited in the HF Heater-Modulated Polar Electrojet

S. P. Kuo

Professor

Department of Electrical Engineering
Polytechnic University

Abstract

A thermal instability responding to the modulation of polar electrojet by HF heater as a potential efficient mechanism for the generation of ELF and VLF waves has been investigated. It is shown that a positive feedback through the electron-neutral collisional heating process can cause the transient response of the plasma to the modulated HF heater to grow exponentially. Considering a sinusoidally modulated HF heater wave having its field amplitude proportional to $|\cos\omega_1 t/2|$, the threshold fields of the thermal instability under normal electrojet conditions are found to be about 0.3 V/m and 0.15 V/m for the o-mode and x-mode operation respectively. For a heater wave field of 0.5 V/m, the instability can be excited by the x-mode heater within a few tenth of a second and by the o-mode heater within one second.

Generation of ELF and VLF Waves by a Thermal Instability Excited in the HF Heater-Modulated Polar Electrojet

S. P. Kuo

INTRODUCTION

The modification of the ionosphere by powerful HF heater waves has been studied experimentally and theoretically for the past thirty years. A large number of unexpected phenomena originating from the region near the HF reflection height in the F region of the ionosphere have been observed besides the originally anticipated electron temperature and density change. Most of these phenomena, though some of them are not yet fully understood, are believed to be attributed to the high frequency plasma turbulence and low frequency density striations excited by the HF heater via various parametric instabilities directly or indirectly. On the other hand, a significant modification effect caused by the HF heater was also expected to occur in the lower ionosphere. Willis and Davis (1973) had pointed out that the perturbations to the ionospheric electrojet current in the D and E regions might give rise to geomagnetic field fluctuations of the order of a few m μ . This prediction was further elaborated by Stubbe and Kopka (1977) as a means of producing artificial geomagnetic micropulsations and ELF and VLF emissions using amplitude-modulated HF heater waves. Experiments to confirm the predictions were first performed with the EISCAT heating facility at Tromso, Norway [Stubbe, et al., 1981]. The report of the successful generation of the ULF radiations by HF heater with modulation of a 10 minute on-off period [Stubbe and Kopka, 1981] stimulated a series of new experiments at the same site as well as at other locations such as Arecibo, Puerto Rico and Fairbanks, Alaska. During the past decade, experiments on the generation of VLF/ELF/ULF radiations in the ionospheric D and E regions by the amplitude-modulated HF heaters have been actively conducted [Stubbe, et al., 1982a, 1982b; Barr and Stubbe, 1984a, 1984b, 1991; Ferraro et al., 1982, 1984; James, et al., 1984; James, 1985; Rietveld et al., 1986; Lee, et al., 1990; McCarrick, et al., 1990; Wong, et al 1990; Barr, et al., 1991]. Encouraging results led to the plan of constructing a new super heating facility in Alaska aimed primarily at the further study of generation of such low frequency radiations [Kossey and Brandt, 1992].

Several processes has been suggested as the potential mechanisms responsible for the generation of VLF/ELF/ULF radiations. They are the modulation of the electrojet current [Stubbe and Kopka, 1977; Fejer and Krenzien, 1982; Papadopoulos, et al, 1985], the thermal filamentation instability of the HF heater (Kuo and Lee, 1983) and the parametric excitation of the Alfvén waves [Papadopoulos, et al., 1982; Ko, et al., 1986] and the whistler waves [Kuo and Lee, 1989], etc. Presently the main focus of the research effort in this subject area is on how to improve the radiation efficiency. It is suggested by Papadopoulos, et al. [1990] that a beam painting approach be able to enlarge the modulated region of the electrojet current and hence enhance the ionospheric antenna gain. As a result, a great deal of design consideration is placed on the sweeping capability of the super heater's antenna system. However, recent experiments by Barr and Stubbe [1991] indicate that in order to take the advantage of the beam painting technology for enhancing the VLF/ELF radiation efficiency, it requires the future super heater to operate at considerably higher power levels than that of the Tromsø facility.

In this paper, we study a thermal instability process for the ELF/VLF generation. It is based on the fact that as the HF heater wave deposits heat into the electrojet plasma, it enhances the electron-neutral collision frequency and the heating rate. If the HF heater is modulated, the heating of the ionosphere is also modulated accordingly. This in turn, can cause the transient temperature response of the plasma to the electrojet current modulation to grow exponentially as an instability. It will be shown in the following that this instability process which depends upon not only the free energy of the HF heater, but also that of the electrojet current has the potential to generate the ELF/VLF radiations more effectively than other known processes.

HF HEATER

In the Tromsø heating experiments for the ELF/VLF generation, the amplitude of the HF heater was modulated with the periodic pulses at the ELF/VLF frequencies. The experiments were operated with variable duty cycles of pulses. The wave field of such an amplitude modulated HF heater can be modeled mathematically as follows:

$$\vec{E}_p = (\hat{x} \pm i\hat{y}) (\epsilon_p/2) \sum_{n=-\infty}^{\infty} P_{T/2}(t - nT_1) e^{i(k_0z - \omega_0t)} + \text{c.c.} \quad (1)$$

where the rectangular pulse $P_{T/2}(t - nT_1)$ is defined to be $P_a(x)=1$ for $|x| < a$, and 0 for $|x| > a$; T_1 is the modulation period and T/T_1 is the percentage of the duty cycle. The upward propagating (in the z direction) heater wave is assumed to be a circularly polarized wave with \pm standing for the R(o) and L(x) modes respectively. The downward magnetic field orients along the $-z$ direction; c.c. represents the complex conjugate. After taking the Fourier series expansion, the amplitude modulation function becomes

$$\sum_{n=-\infty}^{\infty} P_{T/2}(t - nT_1) = T/T_1 + \sum_{n=1}^{\infty} 2 [\sin(n\omega_1 T/2) / n\pi] \cos n\omega_1 t \quad (2)$$

It shows that only a small fraction of the total modulation is contributed to the modulation at the fundamental frequency ω_1 . Its percentage can be increased, however, by introducing a different approach of amplitude modulation. Instead of operating the heater's entire antenna system as a whole, it is separated into, for example, two sets of antenna array radiating at the same power but at slightly different frequencies. Therefore, the heater power delivered to the ionosphere varies temporally at the difference frequency of the two antenna arrays and the resultant wave field of the HF heater can be expressed as

$$\vec{E}_p = (\hat{x} \pm i\hat{y}) (\epsilon_p/2\sqrt{2}) [1 + e^{-i(\omega_1 t + \phi)}] e^{i(k_0z - \omega_0t)} + \text{c.c.} \quad (3)$$

where ϕ is an arbitrary phase factor.

Comparing (3) to (1), the fractional power modulation at the fundamental frequency ω_1 is increased by a factor $\pi/2 \sin(\omega_1 T/2)$. Hence, the modulated fields (3) are expected to cause the modulation of the electrojet current at frequency ω_1 more effectively than the fields (1). In the following analysis, fields (3) are thus adopted as a better approach for ELF/VLF generation. Such an amplitude modulated heater wave field is illustrated in Figure 1.

From the electron momentum equation, the linear velocity responses of electrons to the fields (3) are determined to be

$$\vec{v}_{pe} = -i (\hat{x} \pm i\hat{y}) [e\epsilon_p / 2\sqrt{2}m(\omega_0 \pm \Omega_e + i\nu_{en})] [1 + e^{-i(\omega_1 t + \phi)}] e^{i(k_0 z - \omega_0 t)} + \text{c.c.} \quad (4)$$

where ν_{en} and Ω_e are the electron-neutral collision frequency and the electron cyclotron frequency, respectively.

BACKGROUND CURRENTS

We investigate the modulation of the background electrojet current by the incident HF heater wave. As depicted in Figure 2, $\vec{E}_0 = \hat{x} E_0$ is the dc driving field of the electrojet current and $\vec{B}_0 = -\hat{z} B_0$ is the geomagnetic field. Hence, the background Pederson (x component) and Hall (y component) current densities are given respectively by [Kuo and Lee, 1988]

$$J_x = (n_0 e^2 / m) [\nu_{en} / (\nu_{en}^2 + \Omega_e^2) + (m/M) \nu_i / (\nu_i^2 + \Omega_i^2)] E_0 = -en_0 u_x \quad (5)$$

and

$$J_y = -(n_0 e^2 / m) [\Omega_e / (\nu_{en}^2 + \Omega_e^2) - (m/M) \Omega_i / (\nu_i^2 + \Omega_i^2)] E_0 = -en_0 u_y \quad (6)$$

where ν_i and Ω_i represent the ion collision and cyclotron frequencies, respectively.

In addition, an ac current driven by the wave field of the heater is also present in the background plasma. The current density is given by

$$\vec{J}_{ac} = -en_0 \vec{v}_{pe} \quad (7)$$

Thus the total ohmic heat source in the background plasma is found to be

$$Q = \langle J^2 / \sigma_1^* \rangle \equiv \nu_{en} n_0 m (u^2 + \langle |\vec{v}_{pe}|^2 \rangle) \quad (8)$$

where J is the total current density in the background plasma; $\langle \rangle$ stands for the time average over the heater wave period;

$$\sigma_1^* = (1/\sigma + M\Omega_i^2/n_0e^2v_i)^{-1} = \sigma = n_0e^2/mv_{en};$$

$$u \equiv eE_0/m (v_{en}^2 + \Omega_c^2)^{1/2}; \quad v_{en}v_i \ll \Omega_c\Omega_i \text{ is assumed};$$

$$\langle |\vec{v}_{pe}|^2 \rangle = v_{q\pm}^2 [1 + \cos(\omega_1 t + \phi)] \text{ and } v_{q\pm}^2 = (eE_p/m)^2 / [(\omega_0 \pm \Omega_e)^2 + v_{en}^2].$$

The heat source (8) oscillates at the modulating frequency of the HF heater. It agitates a thermal fluctuation at this frequency in the electron temperature of the electrojet. In terms of the average (over the modulation period) temperature $T_{e0} \equiv \langle T_e \rangle$, the total electron temperature T_e is expressed as $T_e = T_{e0} + \delta T_e$, where δT_e is the temperature perturbation oscillating at the modulating frequency and its harmonics. Since the electron neutral collision frequency v_{en} is proportional to the square root of the electron temperature, it is perturbed in the presence of the thermal fluctuation and can be expanded up to the lowest order temperature perturbation as $v_{en} \equiv v_{e0} [1 + (1/2)(\delta T_e/T_{e0})]$. Consistently, the heat source (8) is also expanded as $Q = Q_0 + \delta Q$ with

$$Q_0 = n_0(v_{e0}/v_{te}^2) \{ (u_0^2 + v_{q0\pm}^2) T_{e0} + \frac{1}{2} v_{q0\pm}^2 \langle \delta T_e \cos(\omega_1 t + \phi) \rangle \} \quad (9)$$

$$\delta Q = \frac{1}{2} n_0(v_{e0}/v_{te}^2) \{ (u_0^2 + v_{q0\pm}^2) \delta T_e + v_{q0\pm}^2 [\delta T_e \cos(\omega_1 t + \phi) - \langle \delta T_e \cos(\omega_1 t + \phi) \rangle] + 2v_{q0\pm}^2 T_{e0} \cos(\omega_1 t + \phi) \} \quad (10)$$

where u_0 and $v_{q0\pm}$ are the corresponding expressions for u and $v_{q\pm}$ with $v_{en} = v_{e0}$; $v_{te}^2 = T_{e0}/m$ and $\langle \rangle$ stands for the time average over the modulation period.

THERMAL INSTABILITY

The basic equation for studying the thermal instability is the electron thermal diffusion equation. It is given by Braginskii [1965]:

$$\frac{3}{2} n_0 \frac{\partial}{\partial t} T_e + P_e \nabla \cdot \vec{v}_e = (R_{//} \nabla_{//}^2 + R_{\perp} \nabla_{\perp}^2) T_e - 3 (m/M_n) n_0 v_{en} (T_e - T_n) + Q \quad (11)$$

where P_e is the electron thermal pressure; M_n and T_n are the mass and temperature of the background neutrals; $R_{//} = 3n_0T_e/2mv_{en}$ and $R_{\perp} = (v_{en}/\Omega_e)^2 R_{//}$ are the thermal diffusion coefficients along and across the geomagnetic field, respectively.

Since the modulated heater wave covers an altitude range much larger than the thickness of the electrojet layer, the heating by the HF heater in the electrojet region may be assumed to be uniform and its diffusion loss along and across the geomagnetic field may be neglected. Since the HF heater is modulated in time only, the associated thermal fluctuation is then expected to vary only in time. Thus the divergent term on the left hand side of (11) can also be neglected in the following analysis.

In terms of the parametric expansions introduced in the previous section, (11) can be decomposed into two, one for T_{e0} and the other one for δT_e . The one for the oscillating fluctuation δT_e is obtained to be

$$\left(\frac{d}{dt} + [a_{\pm} - b_{\pm}\cos(\omega_1 t + \phi)]\right)\delta T_e + b_{\pm}\langle\delta T_e\cos(\omega_1 t + \phi)\rangle = 2b_{\pm}T_{e0}\cos(\omega_1 t + \phi) \quad (12)$$

where $a_{\pm} = v_{e0}[(m/M_n)(3 - T_n/T_{e0}) - (u_0^2 + v_{q0\pm}^2)/3v_{te}^2]$ and $b_{\pm} = (v_{e0}/3)(v_{q0\pm}/v_{te})^2$.

Introducing $\delta T_e = \delta\tilde{T}_e e^{-a_{\pm}t} e^{\beta\sin(\omega_1 t + \phi)}$ into (12), where $\beta_{\pm} = b_{\pm}/\omega_1$, it leads to

$$\frac{d}{dt}\delta\tilde{T}_e = b_{\pm} e^{-\beta\sin(\omega_1 t + \phi)} [2 e^{a_{\pm}t} T_{e0}\cos(\omega_1 t + \phi) - \langle\delta\tilde{T}_e e^{\beta\sin(\omega_1 t + \phi)}\cos(\omega_1 t + \phi)\rangle] \quad (13)$$

Equation (13) can be solved by iteration. This is done by first setting the second term on the right hand side (RHS) of (13) equal to zero for $\delta\tilde{T}_e^{(1)}$. Then substituting the obtained $\delta\tilde{T}_e^{(1)}$ into the second term on the RHS of (13) for evaluating $\delta\tilde{T}_e^{(2)}$. The procedure can be repeated until the desired accuracy of $\delta\tilde{T}_e$ is obtained. However, in the following analysis, only $\delta\tilde{T}_e^{(1)}$ is evaluated because the second term on the RHS of (13) is small compared to the first term. With the aid of the identity

$$e^{-\beta_{\pm} \sin(\omega_1 t + \phi)} = \sum_{l=-\infty}^{\infty} (i)^l I_l(\beta_{\pm}) e^{il(\omega_1 t + \phi)}$$

where I_l is the modified Bessel function of order l . An integration of (13) leads to

$$\delta \tilde{T}_e^{(1)} = \delta \tilde{T}_{e0} + b_{\pm} T_{e0} \sum_{l=-\infty}^{\infty} (i)^{l-1} \frac{(2l/\beta_{\pm}) I_l(\beta_{\pm})}{(a_{\pm} + il\omega_1)} e^{il\phi} (e^{(a_{\pm} + il\omega_1)t} - 1) \quad (14)$$

Therefore, the temperature fluctuation at the fundamental modulation frequency is found to be

$$\begin{aligned} \delta T_{e1} &= -2b_{\pm} T_{e0} \sum_{l=-\infty}^{\infty} (-1)^l \frac{(2l/\beta_{\pm}) I_l(\beta_{\pm})}{(a_{\pm} + il\omega_1)} [(1/\beta_{\pm}) I_l(\beta_{\pm}) \cos(\omega_1 t + \phi) + i I_l(\beta_{\pm}) \sin(\omega_1 t + \phi)] \\ &\quad + 2I_l(\beta_{\pm}) [\delta \tilde{T}_{e0} - b_{\pm} T_{e0} \sum_{l=-\infty}^{\infty} (i)^{l-1} \frac{(2l/\beta_{\pm}) I_l(\beta_{\pm})}{(a_{\pm} + il\omega_1)} e^{il\phi}] e^{-a_{\pm} t} \sin(\omega_1 t + \phi) \\ &= -2b_{\pm} T_{e0} \sum_{l=-\infty}^{\infty} (-1)^l \frac{(2l/\beta_{\pm}) I_l(\beta_{\pm})}{(a_{\pm} + il\omega_1)} \{ [(1/\beta_{\pm}) I_l(\beta_{\pm}) \cos(\omega_1 t + \phi) + i I_l(\beta_{\pm}) \sin(\omega_1 t + \phi)] \\ &\quad - i [I_l(\beta_{\pm})/I_0(\beta_{\pm})] I_l(\beta_{\pm}) e^{-a_{\pm} t} \sin(\omega_1 t + \phi) \} \\ &\cong 8 b_{\pm} T_{e0} [I_1^2(\beta_{\pm}) / \beta_{\pm}^2 (a_{\pm}^2 + \omega_1^2)^{1/2}] [\cos(\omega_1 t + \phi - \theta_{\pm})] - [\beta_{\pm} I_1(\beta_{\pm}) / I_0(\beta_{\pm}) (a_{\pm}^2 + \omega_1^2)^{1/2}] e^{-a_{\pm} t} \sin(\omega_1 t + \phi) \} \\ &= \delta T_{e1s\pm} \cos(\omega_1 t + \phi - \theta_{\pm}) - \delta T_{e1t\pm} e^{-a_{\pm} t} \sin(\omega_1 t + \phi) \end{aligned} \quad (15)$$

where the relationship $\delta \tilde{T}_e - b_{\pm} T_{e0} \sum_{l=-\infty}^{\infty} (i)^{l-1} \frac{(2l/\beta_{\pm}) I_l(\beta_{\pm})}{(a_{\pm} + il\omega_1)} e^{il\phi} = ib_{\pm} T_{e0} \sum_{l=-\infty}^{\infty} (-1)^l \frac{(2l/\beta_{\pm}) I_l^2(\beta_{\pm})}{(a_{\pm} + il\omega_1)}$

derived from $\langle \delta T_e \rangle_{t=0} = 0$ is used and $\theta_{\pm} = \tan^{-1} (\omega_1/a_{\pm})$.

The first term on the RHS of (15) is the steady state response of the plasma to the HF input and its amplitude $\delta T_{e1s\pm}$ is constant in time and is directly proportional to the intensity of the pump (i.e., HF input). On the other hand, the nature of the transient response represented by the second term on the RHS of (15) depends also

on the free energy available in the background plasma. Its amplitude varies in time exponentially. As shown, when $a_{\pm} < 0$, the transient response becomes unstable and the amplitude $\delta T_{e1\pm} e^{-a_{\pm}t}$ of the temperature fluctuation grows in time exponentially as an instability. Since the transient response is triggered by the HF driver, it appears even in the absence of the background noise (i.e. $\delta \bar{T}_e = 0$)

δT_{e1} modulates v_{en} , which, in turn, modulates the electrojet current. From (5) and (6), the perturbed electrojet current density can be expressed as

$$\delta \vec{J}_1 = [v_{e0} \Omega_e^2 \omega_p^2 E_0 / 8\pi(v_{en}^2 + \Omega_e^2)] [\hat{x} + (2v_{e0}/\Omega_e)\hat{y}] \delta T_{e1}/T_{e0} \quad (16)$$

This current density then becomes the source of the ELF/VLF radiation. In general, the amplitude of the radiation $|\delta \vec{E}|$ is proportional to the amplitude of the current source, i.e., $|\delta \vec{E}| \sim |\delta \vec{J}_1|$. Since $|\delta \vec{J}_1| \sim |\delta T_{e1}|$, $|\delta \vec{E}|$ also grows in time exponentially with the onset of the thermal instability.

ESTIMATION OF THE GROWTH RATES

For the quantitative analysis of the concerned instability process, the following E region parameters are used: $T_n \equiv T_i = 300^\circ\text{K}$, $T_{e0} = 1500^\circ\text{K}$, $v_{en} = 5 \times 10^4 \text{sec}^{-1}$, $M_n(\text{NO}^+)/m = 5.52 \times 10^4$, $\Omega_e = 1.35 \text{ MHz}$, and $v_{te} = 1.5 \times 10^5 \text{m/s}$. The heater's frequency employed in the Tromso's experiments is $\omega_0 / 2\pi = 4.04 \text{ MHz}$. In terms of the heater wave field amplitude $\epsilon_p = \epsilon_{p0}$ (V/m) and the dc driving field $E_0 = \epsilon_0$ (mV/m) of the electrojet current, the electron quiver speed can be expressed as $v_{q0\pm} = 6.94 \times 10^3 \epsilon_0 / (1 \pm 1/3)$ (m/s) and $u_0 = 20.75 E_0$ (m/s) respectively. Using these quantities, the growth rates $-a_{\pm}$ of the instability can be written as

$$-a_{\pm} = 2.53 [14.1 \epsilon_{p0}^2 / (1 \pm 1/3)^2 + 1.25 \times 10^{-4} \epsilon_0^2 - 1] \text{ sec}^{-1} \quad (17)$$

Considering the normal situation, $\epsilon_0 = 50$, i.e., $E_0 = 50 \text{ mV/m}$, it requires that $\epsilon_{p0\pm} > 0.22(1 \pm 1/3)$ for a positive growth rate. Thus, the instability thresholds are found to be $\epsilon_{pth\pm} = 0.22(1 \pm 1/3) \text{ V/m}$. The threshold field of the x-mode (-) heater wave is a half of that of the o-mode (+) heater wave. In either case, the threshold

field should be able to be exceeded when the heater is operated with two sets of antenna array radiating in full power and in slightly different frequencies. Without the swelling effect, the wave field of the heater in the E region is about 0.5 V/m. Hence, the growth rates of the instability are calculated to be $\gamma_{\pm} = 3.3$ and 18.3 sec⁻¹ for o-mode and x-mode heater, respectively. It appears that the instability can be excited by the x-mode heater wave within a few tenth of a second in the Tromso's heating experiments. It takes more than five times longer for the instability to grow when the o-mode heater wave is employed.

SUMMARY AND CONCLUSION

A thermal instability has been investigated as an effective mechanism producing the ELF/VLF waves in the polar ionosphere E region where the electrojet current is perturbed by an amplitude-modulated HF heater. It is suggested that the most effective way to modulate the HF heater for the purpose of low frequency wave generation is to operate the heater as two separate sets of antenna array. Each one radiates at the same power but at slightly different frequencies. The frequency difference is determined by the desired ELF/VLF. The two antenna beams are superimposed in the ionosphere to produce a sinusoidally modulated heating power which maximizes the fundamental frequency component and minimizes the harmonics. A positive feedback in response to the modulated heating of the HF pump arises from the electron-neutral collisional heating process which, as shown earlier, can cause the transient temperature response of the electrojet plasma to grow exponentially at the expense of the free energy of the HF heater and the background electrojet current. The thresholds of this instability can be exceeded by the available operational power of the Tromso heating facility. It is expected that ELF/VLF radiation can be generated faster by the x-mode than by the o-mode heater wave by more than five times under the same operational conditions. In the future experiments as the heater power is further increased, the instability mechanism for the generation of ELF and VLF waves becomes even more effective. It is indicated by the numerical results shown in Figs. 3-5 which present the dependence of the growth rate γ_{\pm} , the amplitude of the steady state response $\delta T_{elst\pm}$, and the initial amplitude of the transient response $\delta T_{clt\pm}$ on field amplitude ϵ_{p0} of the heater wave.

ACKNOWLEDGEMENTS

This work was performed at the Geophysics Directorate of the U.S. Air Force's Phillips Laboratory (PL/GP), Hanscom AFB, Massachusetts. The author would like to thank PL for its hospitality. He also wishes to acknowledge the frequent and useful discussions with Dr. Paul Kossey, Mr. John Rasmussen, and Mr. John Heckscher of the PL/GP, and with Professor M.C. Lee of the Massachusetts Institute of Technology.

This work was sponsored by the summer faculty research program of the Air Force office of Scientific Research (AFOSR) Bolling AFB, Washington through the Research and Development Laboratory, Culver City, California.

REFERENCES

- Barr, R. and P. Stubbe, "The Polar Electrojet Antennas as a Source of ELF Radiation in the Earth-Ionosphere Waveguide", *J. Atmos. Terr. Phys.*, 46, 315, 1984a.
- Barr, R. and P. Stubbe, "ELF and VLF Radiation from the Polar Electrojet Antenna", *Radio Sci.*, 19, 1111, 1984b.
- Barr, R., P. Stubbe and H. Kopka, "Long Range Detection of VLF Radiation Produced by Heating the Auroral Electrojet", *Radio Sci.*, 26, 871, 1991.
- Barr, R. and P. Stubbe, "ELF Radiation from the Tromso Super Heater Facility", *Geophys. Res. Lett.*, 18, 1035, 1991.
- Braganskii S. I., "Transport Process in a Plasma" in *Review of Plasma Physics*, Vol. I, edited by M.A. Leontovich, pp. 205, Consultant's Bureau, New York, 1965.
- Fejer, J.A. and E. Krenzien, "Theory of Generation of ULF Pulsations by Ionosphere Modification Experiments", *J. Atmos. Terr. Phys.*, 44, 1075, 1982.
- Ferraro, A.J., H.S. Lee, R. Allshouse, K. Carroll, A.A. Tomko, F.J. Kelly, and R.G. Joiner, "VLF/ ELF radiation from the Ionospheric Dynamo Current System Modulated by Powerful HF Signals", *J. Atmos. Terr. Phys.*, 44, 1113, 1982.
- Ferraro, A.J., H.S. Lee, R. Allshouse, K. Carroll, R. Lunnen, and T. Collins, "Characteristics of Ionospheric ELF Radiation Generated by HF Heating", *J. Atmos. Terr. Phys.*, 46, 855, 1984.
- James, H.G., R.L. Dowden, M.T. Rietveld, P. Stubbe, and H. Kopka, "Simultaneous Observations of ELF Waves from an Artificially Modulated Aurora Electrojet in Space and On the Ground", *J. Geophys. Res.*, 89, 1655, 1984.
- James, H.G., "The ELF Spectrum of Artificially Modulated D/E Region Conductivity", *J. Atmos. Terr. Phys.*, 47, 1129, 1985.

- Ko, K., C.R. Menyuk, A. Reiman, V. Tripathi, P. Palmadesso, and K. Papadopoulos, "ELF Generation in Lower Ionosphere via Collisional Parametric Decay", *J. Geophys. Res.* 91, 10097, 1986.
- Kossey, P.A. and R. B. Brandt, "High Frequency Active Auroral Research Program (HAARP)", presented at National Radio Science Meeting, University of Colorado, Boulder, Colorado, January 7-10, 1992.
- Kuo, S.P. and M.C. Lee, "Earth Magnetic Field Fluctuations Produced by Filamentation Instabilities of Electromagnetic Heater Waves", *Geophys. Res. Lett.*, 10, 979, 1983.
- Kuo, S.P. and M.C. Lee, "Thermal Filamentation Instability Driven by the Auroral Electrojet Current", *J. Geophys. Res.*, 93, 265, 1988.
- Kuo, S.P. and M.C. Lee, "Parametric Excitation of Whistler Waves by HF Heater", *J. Atmos. Terr. Phys.*, 51, 727, 1989.
- Lee, H.S., A. J. Ferraro, and J. V. Olson, "Detection and Characterization of Geomagnetic Pulsations Using HF Ionospheric Heating", *Radio Sci.*, 25, 1429, 1990.
- McCarrick, M.J., D.D. Sentman, A.Y. Wong, R.F. Wuerker, and B. Chouinard, "Excitation of ELF Waves in the Schumann Resonance Range by Modulated HF Heating of the Polar Electrojet", *Radio Sci.* 25, 1291, 1990.
- Papadopoulos, K., R. Sharma, and V. Tripathi, "ELF Generation by Parametric Excitation of Alfvén Waves in the Ionosphere", *J. Geophys. Res.*, 87, 1491, 1982.
- Papadopoulos, K., and C.L. Chang, "Generation of ELF/ULF waves in the ionosphere by Dynamo Processes", *Geophys. Res. Lett.*, 12, 279, 1985.
- Papadopoulos, K., C.L. Chang, P. Vitello, and A. Drobot, "On the Efficiency of Ionospheric ELF Generation", *Radio Sci.*, 25, 1311, 1990.

- Rietveld, M.T., H. Kopka, and P. Stubbe, "D-region Characteristics Deduced from Pulsed Ionospheric Heating Under Auroral Electrojet Conditions", *J. Atmos. Terr. Phys.* 48, 311, 1986.
- Stubbe, P. and H. Kopka, "Modulation of the Polar Electrojet by Powerful HF Waves", *J. Geophys. Res.*, 82, 2319, 1977.
- Stubbe, P., H. Kopka, and R.L. Downen, "Generation of ELF and VLF Waves by Polar Electrojet Modulation: Experimental Results", *J. Geophys. Res.*, 86, 9073, 1981.
- Stubbe, P. and H. Kopka, "Generation of Pc5 Pulsations by Polar Electrojet Modulation: First Experimental Evidence", *J. Geophys. Res.*, 86, 1606, 1981.
- Stubbe, P., H. Kopka, H. Lauche, M.T. Rietveld, A. Brebbe, O. Holt, and R.L. Downen, "Ionospheric Modification Experiment in Northern Scandinavia", *J. Atmos. Terr. Phys.*, 44, 1025, 1982a.
- Stubbe, P., H. Kopka, M.T. Rietveld, and R.L. Downen, "ELF and VLF Wave Generation by Modulated HF Heating of the Current Carrying Lower Ionosphere", *J. Atmos. Terr. Phys.*, 44, 1173, 1982b.
- Willis, J.W., and J.R. Davis, "Radio Frequency Heating Effects on Electron Density in the Lower E Region", *J. Geophys. Res.*, 78, 5710, 1973.
- Wong, A. Y., J. Carroll, R. Dickman, W. Harrison, W. Huhn, B. Lum, M. McCarrick, J. Santoru, C. Schock, G. Wong, and R.F. Wuerker, "High-Power Radiating Facility at the HLPAS Observatory", *Radio Sci.*, 25, 1269, 1990.

FIGURE CAPTIONS

- Fig. 1. Amplitude modulated HF heater wave. Here $T_1=2\pi/\omega_1$ is the modulation period and the envelope is proportional to $|\cos\omega_1 t/2|$.
- Fig. 2. Geometry of coordinate system, the dc electric field \vec{E}_0 , the electrojet current, and the geomagnetic field \vec{B}_0 .
- Fig. 3. The dependence of the instability growth rates on the field amplitude of the heater wave.
- Fig. 4. The dependence of the amplitude of the steady state temperature fluctuations on the field amplitude of the heater wave for two different modulation frequencies (a) $f_1 = 100$ Hz and (b) $f_2 = 1000$ Hz.
- Fig. 5. The dependence of the initial amplitudes of the transient temperature fluctuations on the field amplitude of the heater wave for two different modulation frequencies (a) $f_1 = 100$ Hz and (b) $f_2 = 1000$ Hz.

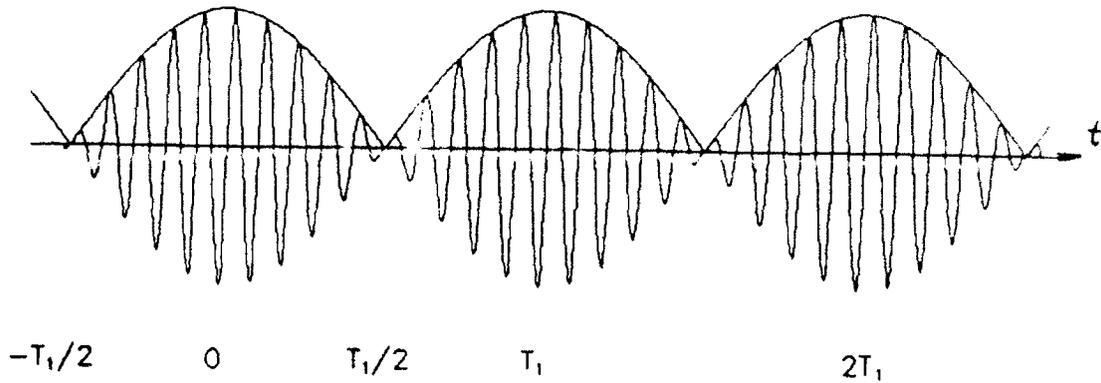


Fig. 1. Amplitude modulated HF heater wave. Here $T_1 = 2\pi/\omega_1$ is the modulation period and the envelope is proportional to $|\cos\omega_1 t/2|$.

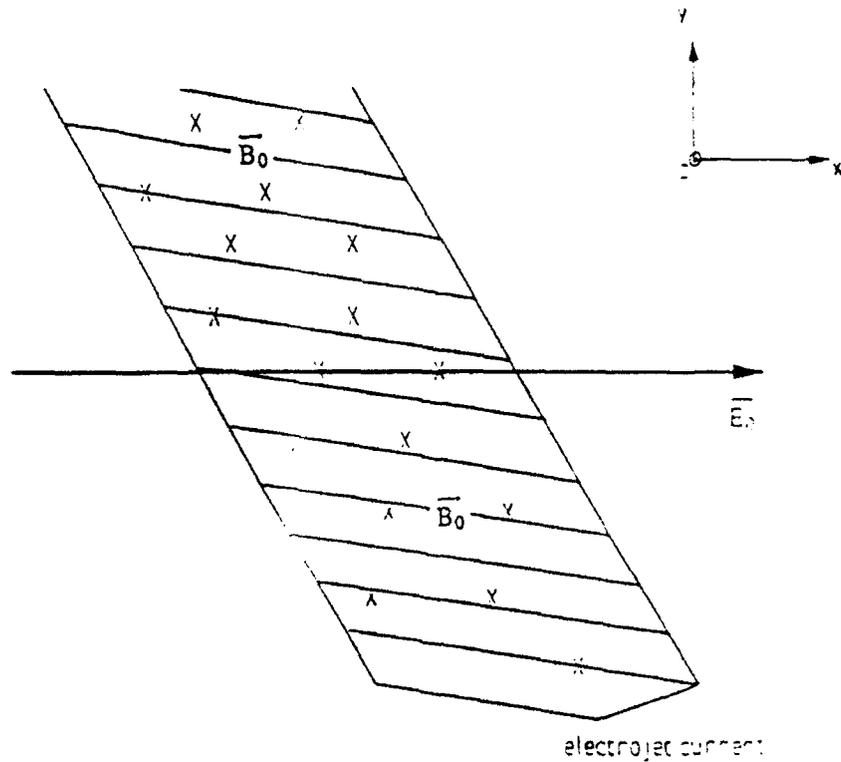


Fig. 2. Geometry of coordinate system, the dc electric field \vec{E}_0 , the electrojet current, and the geomagnetic field \vec{B}_0 .

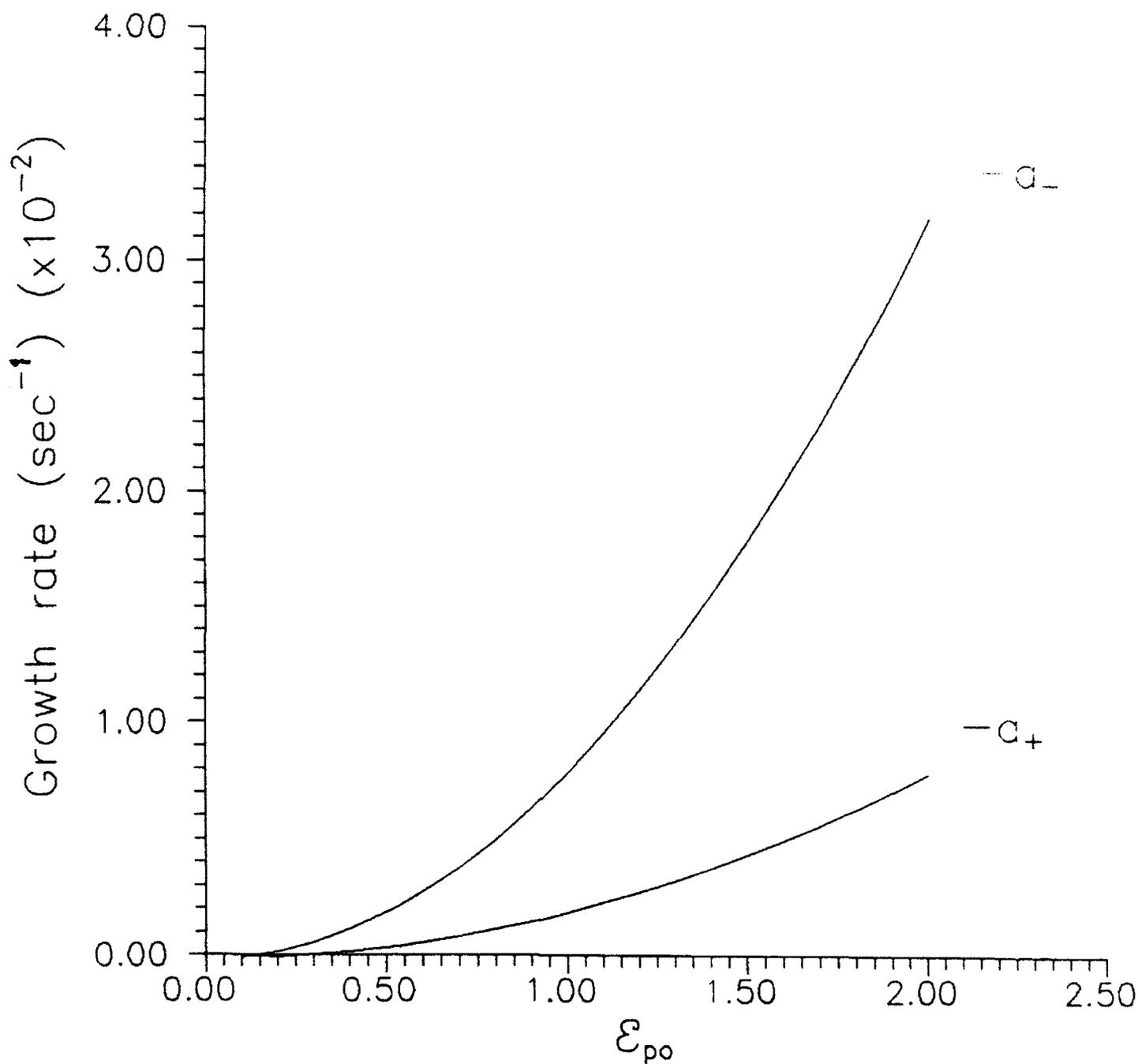


Fig. 3. The dependence of the instability growth rates on the field amplitude of the heater wave.

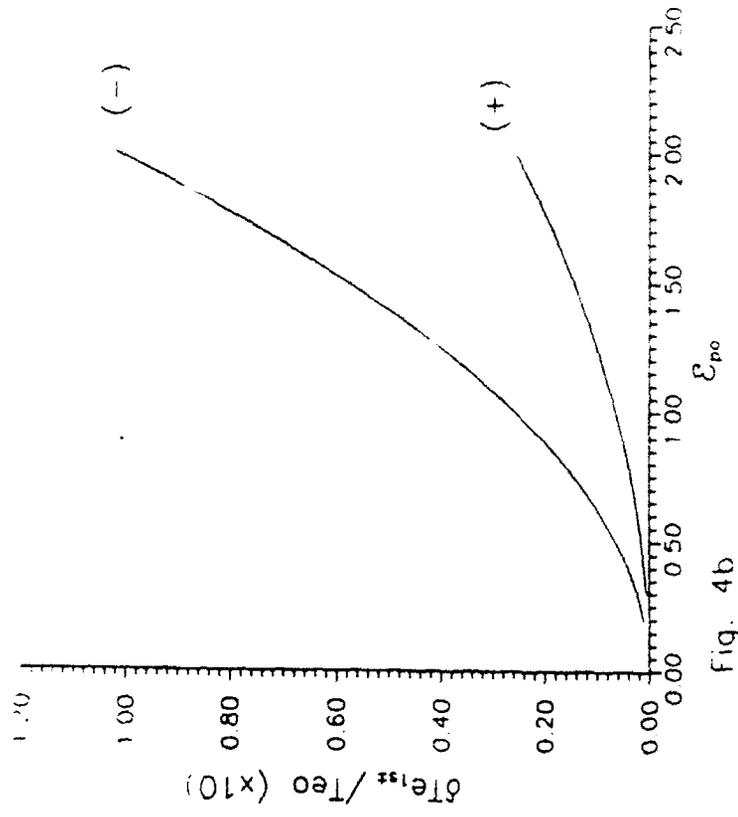


Fig. 4b

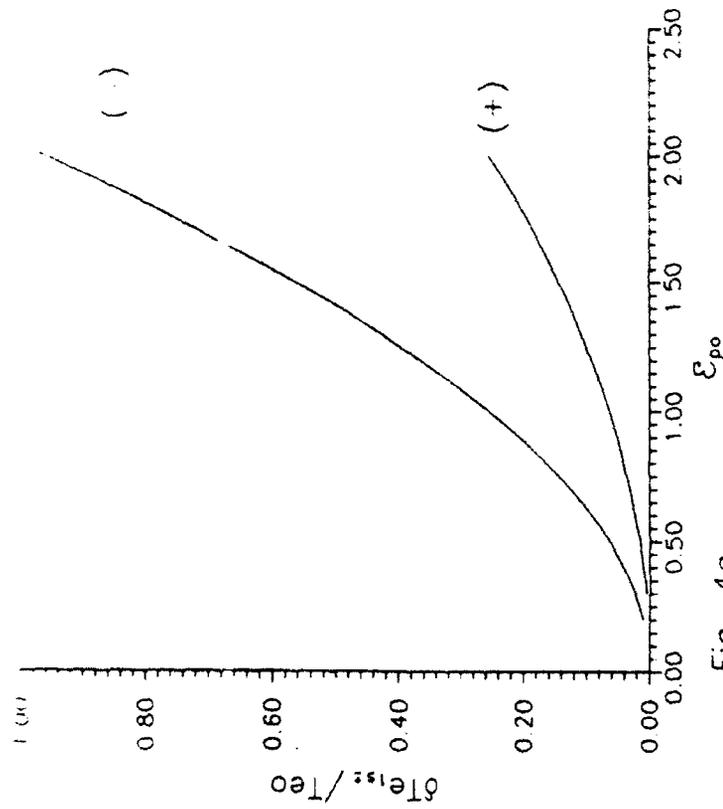


Fig. 4a

Fig. 4. The dependence of the amplitude of the steady state temperature fluctuations on the field amplitude of the heater wave for two different modulation frequencies (a) $f_1 = 100$ Hz and (b) $f_2 = 1000$ Hz.

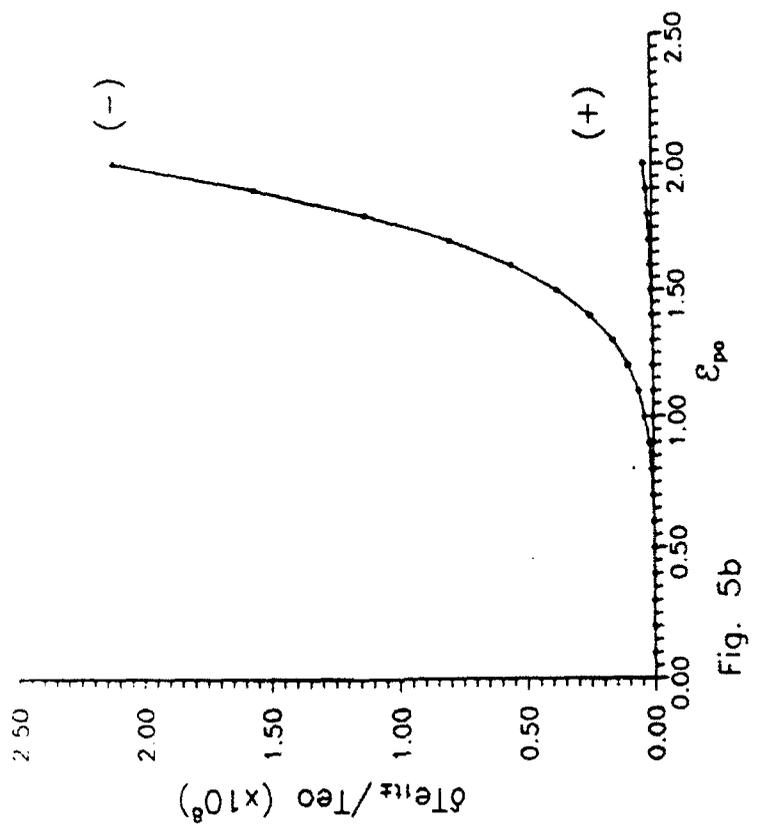


Fig. 5b

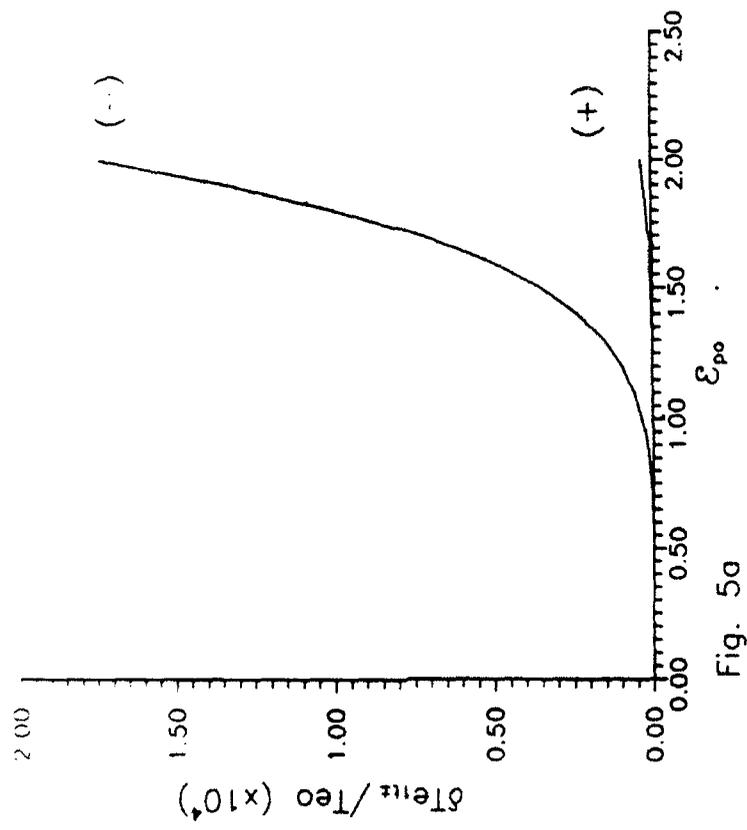


Fig. 5a

Fig. 5. The dependence of the initial amplitudes of the transient temperature fluctuations on the field amplitude of the heater wave for two different modulation frequencies (a) $f_1 = 100$ Hz and (b) $f_2 = 1000$ Hz.

THIS PAGE INTENTIONALLY LEFT BLANK

EQUATION OF RADIATIVE ENERGY TRANSFER
IN ELECTROMAGNETIC THEORY
WITH WAVE-OPTICAL CONTRIBUTIONS

Arvind S. Marathay
Professor
Optical Sciences Center

University of Arizona
Tucson, AZ 85721

Final Report for:
Summer Research Program
Phillips Laboratory
Hanscom AFB

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

September 1992

EQUATION OF RADIATIVE ENERGY TRANSFER
IN ELECTROMAGNETIC THEORY
WITH WAVE-OPTICAL CONTRIBUTIONS

Arvind S. Marathay
Professor
Optical Sciences Center
University of Arizona

Abstract

The conventional equation of radiative transfer is formulated in the frame work of ray optics. It is applicable to scalar waves. Since light is an electromagnetic wave, the ray optic and scalar description is incomplete.

The research reported here formulates the energy transfer problem in the language of wave optics. Starting from Maxwell equations, a radiative energy transfer equation is derived. It shows explicitly the wave optics contributions and includes the source terms generated by induced dielectric polarization. This will provide a basis for studying the effects of polarized light in any state of polarization and properly account for coherence effects in applications.

EQUATION OF RADIATIVE ENERGY TRANSFER
IN ELECTROMAGNETIC THEORY
WITH WAVE-OPTICAL CONTRIBUTIONS

Arvind S. Marathay

(1) INTRODUCTION

The conventional equation of radiative transfer¹ is formulated in the frame work of ray optics. It is applicable to scalar waves.

During the (1992) summer research project, the research effort was directed to obtain an equation of radiative energy transfer in the frame work of electromagnetic theory. In this way effects of the wave nature of polarized light (vector properties of light) can be studied. Furthermore, the source terms generated by induced dielectric polarization are automatically accounted for in the theory. The induced polarization is obtained from the semiclassical matter-field interaction.

The report briefly describes the procedure for obtaining the equation of radiative transfer in wave optics. The literature does not seem to contain such an attempt to account for the wave effects of light in the study of radiative energy transfer.

(2) MAXWELL EQUATIONS AND RAY OPTICS

The Maxwell field equations² in MKSA units are:

$$\begin{aligned} \text{curl } \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} &= 0, \quad \text{div } \mathbf{D} = \rho, \\ \text{curl } \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} &= \mathbf{J}, \quad \text{div } \mathbf{B} = 0 \end{aligned} \tag{1}$$

with the constitutive relations,

$$\mathbf{D} = \epsilon \mathbf{E} = \epsilon_0 \mathbf{E} + \mathbf{P}, \quad \text{and} \quad \mathbf{H} = \frac{1}{\mu_0} \mathbf{B} - \mathbf{M} \tag{2}$$

¹ Chandrasekhar, S. (1960) Radiative Transfer, Dover, New York, NY

²In this document equations in a cluster as in Eq.(1), will be labelled a, b, c, etc.

Specialization to time-harmonic fields,

$$\begin{aligned} \mathbf{E} &= \mathcal{E} \exp(-i\omega t), \\ \mathbf{H} &= \mathcal{H} \exp(-i\omega t) \\ \text{and } \mathbf{P} &= \mathcal{P} \exp(-i\omega t) \end{aligned} \quad (3)$$

where ω is the circular frequency in radian Hertz, and for the case of nonmagnetic media in the absence of free charges and currents, that is,

$$\mathbf{M}=0, \rho=0, \text{ and } \mathbf{J}=0 \quad (4)$$

the field equations become,

$$\begin{aligned} \text{curl } \mathcal{E} &= i\omega\mu_0\mathcal{H}, & \epsilon_0 \text{div } \mathcal{E} &= -\text{div } \mathcal{P} \\ \text{curl } \mathcal{H} &= -i\omega\epsilon_0\mathcal{E} - i\omega\mathcal{P}, & \mu_0 \text{div } \mathcal{H} &= 0 \end{aligned} \quad (5)$$

For regions that are many wavelengths away from the sources, the fields may be expressed in the form³,

$$\begin{aligned} \mathcal{E} &= \mathcal{E}_s \exp(ik_0 S), \\ \mathcal{H} &= \mathcal{H}_s \exp(ik_0 S), \\ \text{and } \mathcal{P} &= \mathcal{P}_s \exp(ik_0 S) \end{aligned} \quad (6)$$

where S is a real function of position \mathbf{r} , not shown explicitly and $k_0 = 2\pi/\lambda_0 = \omega/c$. The fields \mathcal{E}_s , \mathcal{H}_s and \mathcal{P}_s may be complex in general. For states of polarization, such as circular, the functions \mathcal{E}_s , etc have to be complex. The theory is not directly applicable to nonlinear matter-field interactions. However, once the usefulness of the theory is established, the necessary modifications for generality can be carried out at a later time.

With the fields in Eq.(6), Maxwell equations take the form,

³ Born, M. and E. Wolf,(1970) Principles of Optics, fourth edition, section 3.1.1. Pergamon press, Oxford.

$$\begin{aligned}
& \frac{i}{k_0} \nabla \times \mathcal{E}_s + \mathcal{E}_s \times \nabla S + \sqrt{\frac{\mu_0}{\epsilon_0}} \mathcal{H}_s = 0 \\
& \frac{i}{k_0} \nabla \times \mathcal{H}_s + \mathcal{H}_s \times \nabla S - \frac{1}{\sqrt{\epsilon_0 \mu_0}} (\epsilon_0 \mathcal{E}_s + \mathcal{P}_s) = 0 \\
& -\frac{i}{k_0} \nabla \cdot (\epsilon_0 \mathcal{E}_s + \mathcal{P}_s) + (\epsilon_0 \mathcal{E}_s + \mathcal{P}_s) \cdot \nabla S = 0 \\
& -\frac{i}{k_0} \nabla \cdot \mathcal{H}_s + \mathcal{H}_s \cdot \nabla S = 0
\end{aligned} \tag{7}$$

In the so called "ray-optical limit"⁴ where the wavelength λ is negligible compared to any of the dimensions of interest, these equations may be approximated to,

$$\begin{aligned}
& \epsilon_0 \mathcal{E}_s \times \nabla S + \sqrt{\epsilon_0 \mu_0} \mathcal{H}_s = 0 \\
& (\epsilon_0 \mathcal{E}_s + \mathcal{P}_s) \cdot \nabla S = 0 \\
& \sqrt{\epsilon_0 \mu_0} \mathcal{H}_s \times \nabla S - (\epsilon_0 \mathcal{E}_s + \mathcal{P}_s) = 0 \\
& \mathcal{H}_s \cdot \nabla S = 0
\end{aligned} \tag{8}$$

Vectorial multiplication of Eq. (8c) with ∇S and use of Eq. (8a) gives,

$$\sqrt{\epsilon_0 \mu_0} \mathcal{H}_s (|\nabla S|^2 - 1) = \nabla S \times \mathcal{P}_s \tag{9}$$

In a like manner Eq. (8a) leads to,

$$\begin{aligned}
\epsilon_0 \mathcal{E}_s (|\nabla S|^2 - 1) &= \mathcal{P}_s - (\mathcal{P}_s \cdot \nabla S) \nabla S \\
&= (\nabla S \times \mathcal{P}_s) \times \nabla S - \mathcal{P}_s (|\nabla S|^2 - 1)
\end{aligned} \tag{10}$$

When $\mathcal{P}_s = 0$, both Eqs. (9 & 10) lead to the "eikonal" equation of geometrical optics, namely,

⁴Also called the geometrical-optical limit.

$$|\nabla S|^2 = 1 \quad (11)$$

Equations (8) also dictate that in free space, the vectors \mathcal{E} , \mathcal{H} , and ∇S form a right-hand triad of mutually perpendicular vectors. A vector product of Eq.(10) with (9) and use of Eqs.(8) gives a modified eikonal equation,

$$|\nabla S|^2 - 1 = \frac{|\nabla S \times \mathcal{P}_s|}{\sqrt{\epsilon_0 \mu_0} |\mathcal{H}_s|} \quad (12)$$

valid in the medium. In this case the vectors \mathcal{D} , \mathcal{H} and ∇S form a right-hand triad of mutually perpendicular vectors, where $\mathcal{D}_s = \epsilon \mathcal{E}_s - \mathcal{P}_s$ is the familiar \mathcal{D} -field of electromagnetic theory.

For applications of interest to the Phillips Laboratory, it is advantageous to develop the theory in the coordinate system formed by the vectors \mathcal{E}_s , \mathcal{H}_s and ∇S which in the present document will be termed the "free-space" coordinate system.

(3) MODIFIED HELMHOLTZ EQUATION

Return to Eqs.(5a-5d). Take the curl of (5a) and (5c) and use the identity, curl curl = grad div - div grad. It leads to the Helmholtz equations for the \mathcal{E}_s and \mathcal{H}_s respectively,

$$\begin{aligned} \nabla^2 \mathcal{E}_s + k_0^2 \mathcal{E}_s &= -\frac{1}{\epsilon_0} \nabla(\nabla \cdot \mathcal{P}) - \frac{k_0^2}{\epsilon_0} \mathcal{P}, \\ \nabla^2 \mathcal{H}_s + k_0^2 \mathcal{H}_s &= \frac{ik_0}{\sqrt{\epsilon_0 \mu_0}} (\nabla \times \mathcal{P}) \end{aligned} \quad (13)$$

where the dielectric polarization terms on the right-hand side provide the sources. The special form of the fields and the polarization in terms of the eikonal function S of Eqs.(6), is used in the above equations and the resulting equations may be rearranged in the form,

$$\mathbf{K}_a + \frac{i}{k_0} \mathbf{L}_a + \frac{1}{k_0^2} \mathbf{M}_a = 0, \quad a=e \text{ or } h \quad (14)$$

where the subscript, a, equals e or h, according to whether the equation refers to the \mathcal{E}_s field or the \mathcal{H}_s field respectively. The terms \mathbf{K}_a , \mathbf{L}_a , and \mathbf{M}_a are independent of k_0 . These vector terms for the respective cases are:

$$\begin{aligned} \mathbf{K}_e &= \epsilon_0 \mathcal{E}_s (1 - |\nabla S|^2) + \mathcal{P}_s - (\mathcal{P}_s \cdot \nabla S) \nabla S, \\ \mathbf{L}_e &= 2 (\nabla S \cdot \nabla) \epsilon_0 \mathcal{E}_s + (\nabla^2 S) \epsilon_0 \mathcal{E}_s + \nabla (\mathcal{P}_s \cdot \nabla S), \\ \mathbf{M}_e &= \nabla^2 (\epsilon_0 \mathcal{E}_s) + \nabla (\nabla \cdot \mathcal{P}_s) \end{aligned} \quad (15)$$

$$\begin{aligned} \mathbf{K}_h &= \sqrt{\epsilon_0 \mu_0} \mathcal{H}_s (1 - |\nabla S|^2) - \mathcal{P}_s \times \nabla S, \\ \mathbf{L}_h &= 2 \sqrt{\epsilon_0 \mu_0} (\nabla S \cdot \nabla) \mathcal{H}_s + \sqrt{\epsilon_0 \mu_0} (\nabla^2 S) \mathcal{H}_s - \nabla \times \mathcal{P}_s, \\ \mathbf{M}_h &= \sqrt{\epsilon_0 \mu_0} \nabla^2 \mathcal{H}_s \end{aligned} \quad (16)$$

The ray optical limit gives, $\mathbf{K}_a = 0$, for $a = e$ or h . It reproduces Eqs. (9) and (10). The next section uses the free-space coordinate system to discuss the two equations contained in Eq. (14).

(4) EQUATION OF TRANSFER IN WAVE OPTICS

The two equations of Eq. (14) respectively can be rearranged to read,

$$\begin{aligned} & [\nabla^2 + i2k_0 (\nabla S \cdot \nabla) + ik_0 \nabla^2 S] \mathcal{E}_s - k_0^2 (|\nabla S|^2 - 1) \mathcal{E}_s \\ &= -\frac{1}{\epsilon_0} \{ \nabla (\nabla \cdot \mathcal{P}_s) + ik_0 \nabla (\mathcal{P}_s \cdot \nabla S) + k_0^2 [\mathcal{P}_s - (\mathcal{P}_s \cdot \nabla S) \nabla S] \} \end{aligned} \quad (17)$$

$$\begin{aligned} & [\nabla^2 + i2k_0 (\nabla S \cdot \nabla) + ik_0 \nabla^2 S] \mathcal{H}_s - k_0^2 (|\nabla S|^2 - 1) \mathcal{H}_s \\ &= \frac{1}{\sqrt{\epsilon_0 \mu_0}} [+ik_0 (\nabla \times \mathcal{P}_s) - k_0^2 (\nabla S \times \mathcal{P}_s)] \end{aligned} \quad (18)$$

At this point the operators operating on the field variables are transformed with the help of the free-space coordinate system whose axes are labelled with the symbols ξ , η and ζ . The system is oriented in such a way that the ζ axis is parallel to ∇S and the ξ and η axis are aligned parallel to the vectors \mathcal{E}_s , \mathcal{H}_s respectively. In this system, a derivative along the ζ axis is,

$$\frac{\partial}{\partial \zeta} = \frac{\nabla S}{|\nabla S|} \cdot \nabla = \mathbf{e}_s \cdot \nabla \quad (19)$$

For derivatives transverse to the ζ direction, the symbol, ∇_T , used. In this way the gradient operator is expressed by,

$$\nabla = \mathbf{e}_T \nabla_T + \mathbf{e}_s \frac{\partial}{\partial \zeta} = \nabla_T + \mathbf{e}_s \frac{\partial}{\partial \zeta} \quad (20)$$

where \mathbf{e}_T and \mathbf{e}_s are unit vectors with the property, $\mathbf{e}_T \cdot \mathbf{e}_s = 0$. The Laplacian operator takes the form,

$$\nabla^2 = \nabla_T^2 + \frac{\partial^2}{\partial \zeta^2} \quad (21)$$

At this point the parabolic approximation is made. Consider the operator combination appearing in Eqs.(17) and (18),

$$\begin{aligned} & \left(\frac{\partial^2}{\partial \zeta^2} + i2k_0 \frac{\partial}{\partial \zeta} \right) \mathcal{E}_s \\ & = i2k_0 \left(1 - \frac{i\partial}{2k_0 \partial \zeta} \right) \frac{\partial \mathcal{E}_s}{\partial \zeta} \end{aligned} \quad (22)$$

Much like the ray optical approximation, the contribution of the term containing λ_0 is neglected in comparison to unity, that is,

$$\left| \frac{i\lambda_0}{4\pi} \frac{\partial \psi}{\partial \zeta} \right| \ll |\psi| \quad (23)$$

where ψ is some component of the field variable or its derivative. This is referred to as the parabolic approximation. It permits neglecting the second partial derivative on ζ contained in ∇^2 of Eqs. (17) and (18) which take the form,

$$\begin{aligned} & (\nabla_T^2 + i2k_0|\nabla S| \frac{\partial}{\partial \zeta} + ik_0\nabla^2 S - k_0^2(|\nabla S|^2 - 1)) \epsilon_0 \mathcal{E}_s \\ & = -\nabla(\nabla \cdot \mathcal{P}_s) - ik_0\nabla(\mathcal{P}_s \cdot \nabla S) - k_0^2[\mathcal{P}_s - \nabla S(\mathcal{P}_s \cdot \nabla S)] \end{aligned} \quad (24)$$

and

$$\begin{aligned} & [\nabla_T^2 + i2k_0|\nabla S| \frac{\partial}{\partial \zeta} + ik_0\nabla^2 S - k_0^2(|\nabla S|^2 - 1)] \mu_0 \mathcal{H}_s \\ & = \sqrt{\frac{\mu_0}{\epsilon_0}} [ik_0(\nabla \times \mathcal{P}_s) - k_0^2(\nabla S \times \mathcal{P}_s)] \end{aligned} \quad (25)$$

To put these in the form of an energy transfer equation, take a scalar product of Eq.(24) from the left with \mathcal{E}_s^* and add to the result and in turn subtract from the result, the corresponding equation obtained by taking the complex conjugate. The "add" equation is

$$\begin{aligned} & \frac{i}{k_0^2} (\epsilon_0 \mathcal{E}_s^* \nabla_T^2 \mathcal{E}_s - c.c.) + \frac{2}{k_0} |\nabla S| (\epsilon_0 \mathcal{E}_s^* \frac{\partial}{\partial \zeta} \mathcal{E}_s + c.c.) + \frac{2}{k_0} \nabla^2 S \epsilon_0 \mathcal{E}_s^* \mathcal{E}_s \\ & = \frac{i}{k_0^2} [\mathcal{E}_s^* \nabla(\nabla \cdot \mathcal{P}_s) - c.c.] - \frac{1}{k_0} [\mathcal{E}_s^* \nabla(\mathcal{P}_s \cdot \nabla S) + c.c.] \\ & \quad + \{i \mathcal{E}_s^* \cdot [\mathcal{P}_s - (\mathcal{P}_s \cdot \nabla S) \nabla S] + c.c.\} \end{aligned} \quad (26)$$

and the "subtract" equation is

$$\begin{aligned} & \frac{1}{k_0^2} (\epsilon_0 \mathcal{E}_s^* \nabla^2 \mathcal{E}_s + c.c.) + \frac{2}{k_0} |\nabla S| (i \epsilon_0 \mathcal{E}_s^* \frac{\partial}{\partial \zeta} \mathcal{E}_s + c.c.) - 2(|\nabla S|^2 - 1) \epsilon_0 \mathcal{E}_s^* \mathcal{E}_s \\ & = -\frac{1}{k_0^2} [\mathcal{E}_s^* \nabla(\nabla \cdot \mathcal{P}_s) + c.c.] - \frac{1}{k_0} [i \mathcal{E}_s^* \nabla(\mathcal{P}_s \cdot \nabla S) + c.c.] \\ & \quad - \{ \mathcal{E}_s^* \cdot [\mathcal{P}_s - (\mathcal{P}_s \cdot \nabla S) \nabla S] + c.c. \} \end{aligned} \quad (27)$$

The notation c.c. stands for complex conjugate. By use of the transverse and longitudinal derivatives of Eq.(19) through (21), a part of the left side of Eq.(26) can be transformed to,

$$\begin{aligned} & \frac{2}{k_0} |\nabla S| (\epsilon_0 \mathcal{E}_s \cdot \frac{\partial}{\partial \zeta} \mathcal{E}_s^* + c.c.) + \frac{2}{k_0} (\nabla^2 S) \epsilon_0 \mathcal{E}_s \cdot \mathcal{E}_s^* \\ &= \frac{2}{k_0} \nabla_T \cdot (\epsilon_0 \mathcal{E}_s \cdot \mathcal{E}_s^* \nabla S) + \frac{2}{k_0} \frac{\partial}{\partial \zeta} (\epsilon_0 \mathcal{E}_s \cdot \mathcal{E}_s^* |\nabla S|) \end{aligned} \quad (28)$$

Thus the "add" equation becomes,

$$\begin{aligned} & \frac{i}{k_0^2} (\epsilon_0 \mathcal{E}_s \cdot \nabla_T^2 \mathcal{E}_s^* - c.c.) + \frac{2}{k_0} \nabla_T \cdot (\epsilon_0 \mathcal{E}_s \cdot \mathcal{E}_s^* \nabla S) + \frac{2}{k_0} \frac{\partial}{\partial \zeta} (\epsilon_0 \mathcal{E}_s \cdot \mathcal{E}_s^* |\nabla S|) \\ &= \frac{i}{k_0^2} [\mathcal{E}_s^* \cdot \nabla (\nabla \cdot \mathcal{P}_s) - c.c.] - \frac{1}{k_0} [\mathcal{E}_s^* \cdot \nabla (\mathcal{P}_s \cdot \nabla S) + c.c.] \\ & \quad + \{i \mathcal{E}_s^* \cdot [\mathcal{P}_s - (\mathcal{P}_s \cdot \nabla S) \nabla S] + c.c.\} \end{aligned} \quad (29)$$

The eikonal equation is contained in the "subtract" equation (27) in the ray-optical limit. The equation of energy transfer is in the "add" equation (29).

The ray-optical limit of Eq.(27) yields the eikonal,

$$(|\nabla S|^2 - 1) \epsilon_0 \mathcal{E}_s \cdot \mathcal{E}_s^* = \frac{1}{2} \{ \mathcal{E}_s^* \cdot [\mathcal{P}_s - (\mathcal{P}_s \cdot \nabla S) \nabla S] + c.c. \} \quad (30)$$

and the corresponding limit of Eq.(29) is,

$$0 = \{i \mathcal{E}_s^* \cdot [\mathcal{P}_s - (\mathcal{P}_s \cdot \nabla S) \nabla S] + c.c.\} \quad (31)$$

Assuming for the moment that conditions (30) and (31) are fulfilled, then Eq.(29) leads to the radiative energy transfer equation,

$$\begin{aligned}
& \frac{\partial}{\partial \zeta} (\epsilon_0 \mathcal{E}_s \cdot \mathcal{E}_s^* |\nabla S|) + \nabla_{\mathbf{r}} \cdot (\epsilon_0 \mathcal{E}_s \cdot \mathcal{E}_s^* \nabla S) + \frac{i}{2k_0} (\epsilon_0 \mathcal{E}_s \cdot \nabla_{\mathbf{r}}^2 \mathcal{E}_s^* - c.c.) \\
& = -\frac{1}{2} [\mathcal{E}_s^* \cdot \nabla (\mathcal{P}_s \cdot \nabla S) + c.c.] + \frac{i}{2k_0} [\mathcal{E}_s^* \cdot \nabla (\nabla \cdot \mathcal{P}_s) - c.c.]
\end{aligned} \tag{32}$$

whereas Eq. (27) gives,

$$\begin{aligned}
& |\nabla S| (i\epsilon_0 \mathcal{E}_s \cdot \frac{\partial}{\partial \zeta} \mathcal{E}_s + c.c.) + \frac{1}{2k_0} (\epsilon_0 \mathcal{E}_s \cdot \nabla_{\mathbf{r}}^2 \mathcal{E}_s + c.c.) \\
& = -\frac{1}{2} [i\mathcal{E}_s^* \cdot \nabla (\mathcal{P}_s \cdot \nabla S) + c.c.] - \frac{1}{2k_0} [\mathcal{E}_s^* \cdot \nabla (\nabla \cdot \mathcal{P}_s) + c.c.]
\end{aligned} \tag{33}$$

The ray-optical limit of Eq. (32) leads to the conventional equation of transfer which is independent of wavelength. The wave-optical contributions to this equation are contained in the coefficients of $1/k_0$. The above discussion applies only to the electric field part of the total equation of transfer.

The magnetic part is obtained in a similar fashion by working with Eq. (25). Without going into the details the relevant results are listed. First the eikonal equation is,

$$(|\nabla S|^2 - 1) \mu_0 \mathcal{H}_s \cdot \mathcal{H}_s^* = -\frac{1}{2} \sqrt{\frac{\mu_0}{\epsilon_0}} [\mathcal{H}_s^* \cdot (\nabla S \times \mathcal{P}_s) + c.c.] \tag{34}$$

It is found in the ray-optical limit of the "subtract" equation, while in this limit the "add" equation gives,

$$0 = \frac{1}{2} \sqrt{\frac{\mu_0}{\epsilon_0}} i [\mathcal{H}_s^* \cdot (\nabla S \times \mathcal{P}_s) - c.c.] \tag{35}$$

The equivalence of the two eikonal equations (34) and (30) needs to be established but not given in this report. Assuming for the moment that Eqs. (34) and (35) are fulfilled then the "add" equation leads to the energy transfer equation,

$$\begin{aligned}
& \frac{\partial}{\partial \zeta} (\mu_0 \mathcal{H}_s \cdot \mathcal{H}_s^* |\nabla S|) + \nabla_T \cdot (\mu_0 \mathcal{H}_s \cdot \mathcal{H}_s^* \nabla S) + \frac{i}{2k_0} (\mu_0 \mathcal{H}_s \cdot \nabla_T^2 \mathcal{H}_s^* - c.c.) \\
& = \frac{1}{2} \sqrt{\frac{\mu_0}{\epsilon_0}} [\mathcal{H}_s^* \cdot (\nabla \times \mathcal{P}_s) + c.c.]
\end{aligned} \tag{36}$$

while the associated "subtract" equation gives,

$$\begin{aligned}
& |\nabla S| (i \mu_0 \mathcal{H}_s^* \cdot \frac{\partial}{\partial \zeta} \mathcal{H}_s + c.c.) + \frac{1}{2k_0} (\mu_0 \mathcal{H}_s^* \cdot \nabla_T^2 \mathcal{H}_s + c.c.) \\
& = \sqrt{\frac{\mu_0}{\epsilon_0}} i [\mathcal{H}_s^* \cdot (\nabla \times \mathcal{P}_s) - c.c.]
\end{aligned} \tag{37}$$

A ray-optical limit of the transfer equation (36) gives the magnetic part of the conventional equation.

The energy transfer equations for the electric and the magnetic cases, may be combined together to give an equation of the total energy. The time averaged electric, magnetic and total energy densities are respectively defined by the symbols,

$$\begin{aligned}
w_{se} &= \frac{1}{2} \epsilon_0 \mathcal{E}_s \cdot \mathcal{E}_s^* \\
w_{sm} &= \frac{1}{2} \mu_0 \mathcal{H}_s \cdot \mathcal{H}_s^* \\
\text{and } w_s &= w_{se} + w_{sm}
\end{aligned} \tag{38}$$

Addition of Eqs. (32) and (36) gives the radiative transfer equation for the total electric and magnetic energy densities,

$$\begin{aligned}
& \frac{\partial}{\partial \zeta} (w_s |\nabla S|) + \nabla_T \cdot (w_s \nabla S) \\
& + \frac{i}{4k_0} (\epsilon_0 \mathcal{E}_s \cdot \nabla_T^2 \mathcal{E}_s^* - c.c.) + \frac{i}{4k_0} (\mu_0 \mathcal{H}_s \cdot \nabla_T^2 \mathcal{H}_s^* - c.c.) \\
& = -\frac{1}{4} [\mathcal{E}_s^* \cdot \nabla (\mathcal{P}_s \cdot \nabla S) + c.c.] + \frac{1}{4} \sqrt{\frac{\mu_0}{\epsilon_0}} [\mathcal{H}_s^* \cdot (\nabla \times \mathcal{P}_s) + c.c.] \\
& + \frac{i}{4k_0} [\mathcal{E}_s^* \cdot \nabla (\nabla \cdot \mathcal{P}_s) - c.c.]
\end{aligned} \tag{39}$$

The first two terms of the left side are like the conventional equation of transfer, it contains the longitudinal derivative of a quantity proportional to the total energy density, and the transverse divergence of the flow vector whose direction is parallel to the light rays. The last two terms of the left side are the wave-optics contribution which contain the field variables. The right side contains the ray-optic contributions independent of the wavelength. The coefficient of $1/k_0$ on the right-hand side is the wave-optics contribution of the interaction energy.

The derivation of Eq.(39) follows the conventional way of making the parabolic approximation, wherein the second partial derivative on ζ is neglected as discussed in Eqs.(22) and (23). It is important to note however, that in the context of wide angle wave-optics contributions, the $\partial^2 / (\partial\zeta^2)$ term is just as important as the terms containing the transverse Laplacian (∇_{\perp}^2) appearing on the left-hand side of Eq.(39). Thus to study wave-optics contributions the parabolic approximation should not be made. Instead of (∇_{\perp}^2) the Laplacian (∇^2) appears in Eq.(39) which then becomes an exact statement of energy transfer.

(5) Outlook

The research effort during the summer was directed toward obtaining the wave-optics contribution to the equation of energy transfer. It is derived by using Maxwell equations and the resulting Helmholtz equations of the fields. The equation of energy transfer with the wave-optics contribution is new and it reduces correctly to the conventional equation of ray optics when the contributions proportional to the wavelength are neglected.

Wave-theoretic radiometry was not used. Although it should be possible to derive a similar equation by using the spectral radiance function with units of watts/(m² sr Hz). The present work paves the way for the study of the propagation of correlation functions of the electromagnetic field.

ACKNOWLEDGEMENT

Helpful discussions and comments from Dr. R. H. Picard of the Phillips Laboratory and Professor M. J. Beran of the Catholic University, Washington, DC are gratefully acknowledged.

ION-MOLECULE REACTIONS AT HIGH TEMPERATURES

Melani Menendez-Barreto
Graduate Student
Department of Physics
University of Puerto Rico at Mayaguez
Mayaguez, PR 00680

Jeffrey F. Friedman
Assistant Professor
Department of Physics
University of Puerto Rico at Mayaguez
Mayaguez, PR 00680

Thomas M. Miller
Professor
Department of Physics and Astronomy
University of Oklahoma
Norman, OK 73019

Final Report for:
Summer Research Program
Phillips Laboratory, Geophysics Directorate

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, DC

August 1992

ION-MOLECULE REACTIONS AT HIGH TEMPERATURES

Melani Menéndez-Barreto
Graduate Student
Department of Physics
University of Puerto Rico at Mayaguez

Jeffrey F. Friedman
Assistant Professor
Department of Physics
University of Puerto Rico at Mayaguez

Thomas M. Miller
Professor
Department of Physics and Astronomy
University of Oklahoma

Abstract

A flowing afterglow apparatus designed for the measurement of ion-molecule reaction rate coefficients at temperatures higher than any previous work was debugged and put into operation during the summer of 1992. Ion-molecule reaction rate coefficients were measured for a variety of systems in the temperature range 300-1200 K: $O^- + H_2, D_2, N_2, CO, NO,$ and CH_4 ; $Ar^+ + H_2, O_2, CO, NO,$ and CH_4 ; $O_2^+ + CH_4$; and $Cl^- + CH_3Br$ and CH_3I .

ION-MOLECULE REACTIONS AT HIGH TEMPERATURES

Melani Menéndez-Barreto, Jeffrey F. Friedman,
and Thomas M. Miller

INTRODUCTION

It has long been recognized that information on ion-molecule reactions is desirable for both scientific and practical reasons.¹ New reaction channels can open up at high temperatures as rotational and vibrational states are excited in both the ion and target molecule. The somewhat controversial topic of "entropy-driven" reactions could more readily be investigated at high temperatures, as emphasized by Meot-Ner.²

The U.S. Air Force sponsors a great deal of research on low-energy interactions of electrons, ions, atoms, and molecules because of the need to understand and influence plasmas which occur naturally (e.g., auroras and the ionosphere) or man-made (e.g., around transonic aircraft, re-entry vehicles, or engine exhausts). Generally speaking, the ultimate goal is to be able to eliminate free electrons from the plasma to enhance radiowave communications and/or reduce radar signatures. In order to gain control of the plasma, one must model the plasma using perhaps thousands of reaction rate coefficients for the interacting species.

One difficulty is that laboratory studies of ion-molecule reactions are conventionally carried out close to, or within 200-300 celsius degrees, of room temperature, while many of the plasmas of interest to the Air Force are at much higher temperatures (say, 1000-2000 K). Modelers are forced to make educated guesses as to the high temperature behavior of reaction rates and products.

For some reactions, the translational energy dependence of the reaction rate coefficients is known up to energies corresponding to thousands of degrees, from drift tube experiments. However, the

translational energy data are not directly applicable to high temperature plasmas because the target molecules in those experiments were not at equivalently high temperatures, and the projectile ions may or may not be excited by successive collisions with a buffer gas.

A small fraction of the total ion-molecule reaction literature includes measurements of reaction rate coefficients as a function of temperature.³ These measurements typically extend into the 500-600 K range, limited by the type of materials and heaters used in the apparatus. A single experiment has gone beyond this conventional temperature limit, prior to the present work. In 1974, a group at NOAA modified the heaters on a flowing afterglow apparatus and measured reaction rate coefficients for several positive-ion reactions involving atmospheric species, up to 900 K.⁴

We now have measurements as high as 1200 K for at least a dozen ion-molecule systems.

EXPERIMENTAL METHOD

We used a newly-constructed flowing afterglow apparatus at Hanscom AFB that was specially designed for high temperatures. The flowing afterglow technique was developed in the 1960s for the study of ion-molecule reactions by a group at NOAA laboratories (then part of the National Bureau of Standards). The flowing-afterglow method is now a well-established technique used in 20-30 laboratories around the world.

The experimental method involves interacting a swarm of ions with a known number density (n) of reactant molecules for a specific time (t). Exponential attenuation of the primary ion signal (I , which is proportional to the primary ion density) will occur if reactions take place which generate other species:

$$I = I_0 \exp (-k n t),$$

where the signal at $t=0$ is I_0 . The ion-molecule reaction rate coefficient (k) is given by:

$$k = - n^{-1} t^{-1} \ln (I/I_0).$$

The experiment was carried out in a flowing buffer gas to effectively transform the time measurement into one of distance.

The measurement procedure involved recording the primary ion count rate as the neutral reactant concentration was stepped from zero to an amount that affected at least a one decade attenuation in the primary ion signal. The (negative) logarithmic slope of the primary ion count rate vs reactant concentration is equal to $-kt$. The reaction time, t , was readily determined from a measurement of the ion velocity and the known reaction distance.

Mass spectra identified the ionic products of the reaction under study. Electron detachment reactions were revealed by the lack of ionic product. Neutral products could not be observed, but could be inferred from known energetics.

APPARATUS

In the flowing afterglow apparatus, an ion swarm (e.g., consisting of O^- ions) was created by interaction between a source gas (e.g., N_2O) and an electron beam at the upstream end of the flow tube. These primary ions drifted approximately 40 cm toward an inlet for neutral reactants. The primary ions were allowed to mix with the neutrals, reacting for 63 cm before a sample of the final ion swarm was taken through an aperture into a mass spectrometer and detection chamber.

The temperature of the flow tube was measured in these experiments by two chromel-alumel thermocouples in the ion-molecule interaction region, inside the flow tube. The ion velocity in the interaction region was measured by applying a repetitive 70-us electrical pulse to these thermocouples, sequentially, and timing the arrival of the disturbance at the detector. The ion velocity was typically 15,000 cm/s.

The concentration of neutral reactant in the flow tube has been determined by measuring the flow rate of reactant, and comparing to the flow rate and pressure of the buffer gas. The reactant gas concentration was typically 1/1000th of the buffer gas concentration. The buffer gas pressure ranged from 0.3-1.3 torr.

The Phillips Laboratory flowing afterglow apparatus was designed to reach a temperature of 2000 K in 2 (possibly 3) stages of development. The first (present) stage utilizes a stainless steel flow tube wrapped with commercial heating tapes and zirconia insulation, and has proven capable of routine operation at temperatures at least as high as 1200 K. Our high-temperature flowing-afterglow (HTFA) apparatus has a number of specific accommodations for elevated temperatures.

1. Only stainless steel (flow tube and inlet tubes), molybdenum (ion sampling plate), and alumina (electrical insulator) materials are in the hot region of the flow tube. In the final stage of development, a ceramic flow tube will be installed, surrounded by a furnace.
2. All inlet tubes enter the flow tube from the upstream end rather than through radial ports in the side of the flow tube. Thus there is nothing along the walls of the flow tube to interfere with the heaters and insulation.
3. Water cooling coils are soldered around flanges where vacuum seals are made, and around

the mass spectrometer housing.

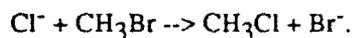
4. The downstream end of the flow tube (the end of the interaction region for ion-molecule reactions) is welded to a thin (1/16") stainless steel plate which has an O.D. much greater than the O.D. of the flow tube (16" vs 3"). The circumference of this plate is water cooled, to allow an o-ring vacuum seal to be made. The thermal conductivity of the thin stainless steel plate is low enough that the end of the flow tube can be maintained as hot as the rest of the interaction region. The only drawback to the thin stainless steel plate is that it bows inward at high temperatures (by 1/8 - 1/4 inch). The buffer gas velocity in a flowing afterglow normally increases steadily with temperature; in our HTFA, the buffer gas velocity (and hence the ion velocity) increased with temperature until 800 K, then decreased with temperature as the end of the flow tube bowed toward the sampling cone, decreasing conductance to the vacuum pump.

5. A large (2'x2'x5') aluminum vacuum box surrounds the flow tube and mass spectrometer housing to provide insulation, and to deal with inevitable leaks from the flow tube once a ceramic flow tube is installed.

Typical data, for the reaction $O^- + H_2$ at 1100 K, are shown in Fig. 1. The logarithmic slope of the attenuation of O^- signal is $-0.1302 \times 10^{-11} \text{ cm}^3$. Combining the known reaction distance (63.7 cm, which includes an 0.87-cm thermal expansion correction) and the measured ion velocity (16,082 cm/s) gives the reaction time, 3.96 ms. It follows from the formulas given above that the reaction rate coefficient for the $O^- + H_2$ reaction is $k = 3.29 \times 10^{-10} \text{ cm}^3/\text{s}$ at 1100 K. The reaction leads primarily to $e^- + H_2O$, with a minor $OH^- + H$ channel evident.

RESULTS AND DISCUSSION

Our first high-temperature data were for the classical nucleophilic displacement reaction,



Previous work in this laboratory on a selected-ion flow-tube (SIFT) apparatus had shown a steady decline in the reaction rate coefficient for this reaction, up to nearly 600 K. As shown in Fig. 2, we found a minimum in the reaction rate coefficient at about 550 K, followed by an increase of almost 2 orders of magnitude, making the apparent rate coefficient essentially collisional at 1200 K.

However, CH_3Br could easily be undergoing heterogeneous chemistry on the hot stainless steel surfaces (yielding, for example, Br_2 or KBr). So the measured reaction rate coefficients at high temperatures may not pertain to CH_3Br . Continuation of this research will involve changes to the HTFA apparatus to permit electron-beam residual-gas analysis, hopefully to settle the question of the identity of neutral reactants at high temperatures. An abstract on this work has been accepted for presentation at the Gaseous Electronics Conference in Boston in October 1992.⁵

Results for $\text{Cl}^- + \text{CH}_3\text{I}$ exhibit similar behavior, though the minimum in the curve occurs near 700 K.

To avoid the dissociation problem, we next used a number of very stable gases; Cl^- will not react with most (Cl^- is very stable itself), so we turned to a more reactive projectile, O^- . Fig. 3 shows results for $\text{O}^- + \text{H}_2$, D_2 , NO , CO , and CH_4 . The hydrogen reaction was discussed in the previous section. The NO and CO reactions lead to electron detachment, with neutral NO_2 and CO_2 formed, respectively. The CH_4 reaction yields $\text{OH}^- + \text{CH}_3$ product.

Figure 4 shows results for less stable gases, NH_3 and N_2O , reacting with O^- . The apparent rate coefficients seem to fall above about 800 K, which we attribute to heterogeneous chemistry on the hot stainless steel walls of the apparatus, yielding non-reactive products. As with the methyl halides, residual

gas analysis would be helpful in understanding these results.

Results for Ar^+ with various gases (H_2 , O_2 , CO , NO , and CH_4) are not shown here because we suspect that the Ar^+ formation had not been allowed to go to completion prior to the neutral reactant inlet. We intend to repeat these measurements, and make diagnostic tests, at some point in the next six months. We also have data for $\text{O}_2^+ + \text{CH}_4$ up to 1100 K. These results compare well with previous SIFT work, and indicate that vibrational excitation is causing the reaction rate coefficient to increase with temperature.

Positive ion reactions studies may prove to be limited in temperature. The walls of the hot flow tube emit positive alkali ions, which swamp the positive ion signal above 1100 K. We have detected K^+ ions at temperatures as low as 600 K. At 1200 K, Na^+ , K^+ , Rb^+ , and Cs^+ were observed with the electron-impact ion source off. (In fact, turning on the ion source made no difference.)

Finally, we carried out a few measurements⁶ at room temperature on proton-transfer reactions involving strong gas-phase acids, at the request of colleagues.

REFERENCES

- ¹See, for example, *Gas Phase Ion Chemistry*, Vol. 2, Ed. M. T. Bowers (Academic Press, New York, 1979).
- ²M. Meot-Ner (Mautner), *J. Phys. Chem.* **95**, 6580 (1991).
- ³Y. Ikezoe, S. Matsuoka, M. Takebe, and A. Viggiano, *Gas Phase Ion-Molecule Reaction Rate Constants Through 1986* (Maruzen Company, Ltd., Tokyo, 1987).
- ⁴W. L. Lindinger, F. C. Fehsenfeld, A. L. Schmeltekopf, and E. E. Ferguson, *J. Geophys. Res.* **79**, 4753 (1974).
- ⁵M. Menendez-Barreto, J. F. Friedman, T. M. Miller, A. A. Viggiano, R. A. Morris, A. E. S. Miller,

J. M. Van Doren, and J. F. Paulson, "Ion-Molecule Reactions at High Temperatures," *Bull. Amer. Phys. Soc.* (to be published).

⁶A. E. S. Miller, T. M. Miller, G. K. Rekha, M. Menendez-Barreto, and J. F. Friedman, "Unusual Kinetics in Gas-Phase Proton Transfer to $\text{Mn}(\text{CO})_5^-$," (to be submitted to *J. Amer. Chem. Soc.*).

Fig. 1. Typical data. O- + H2 at 1100 K.

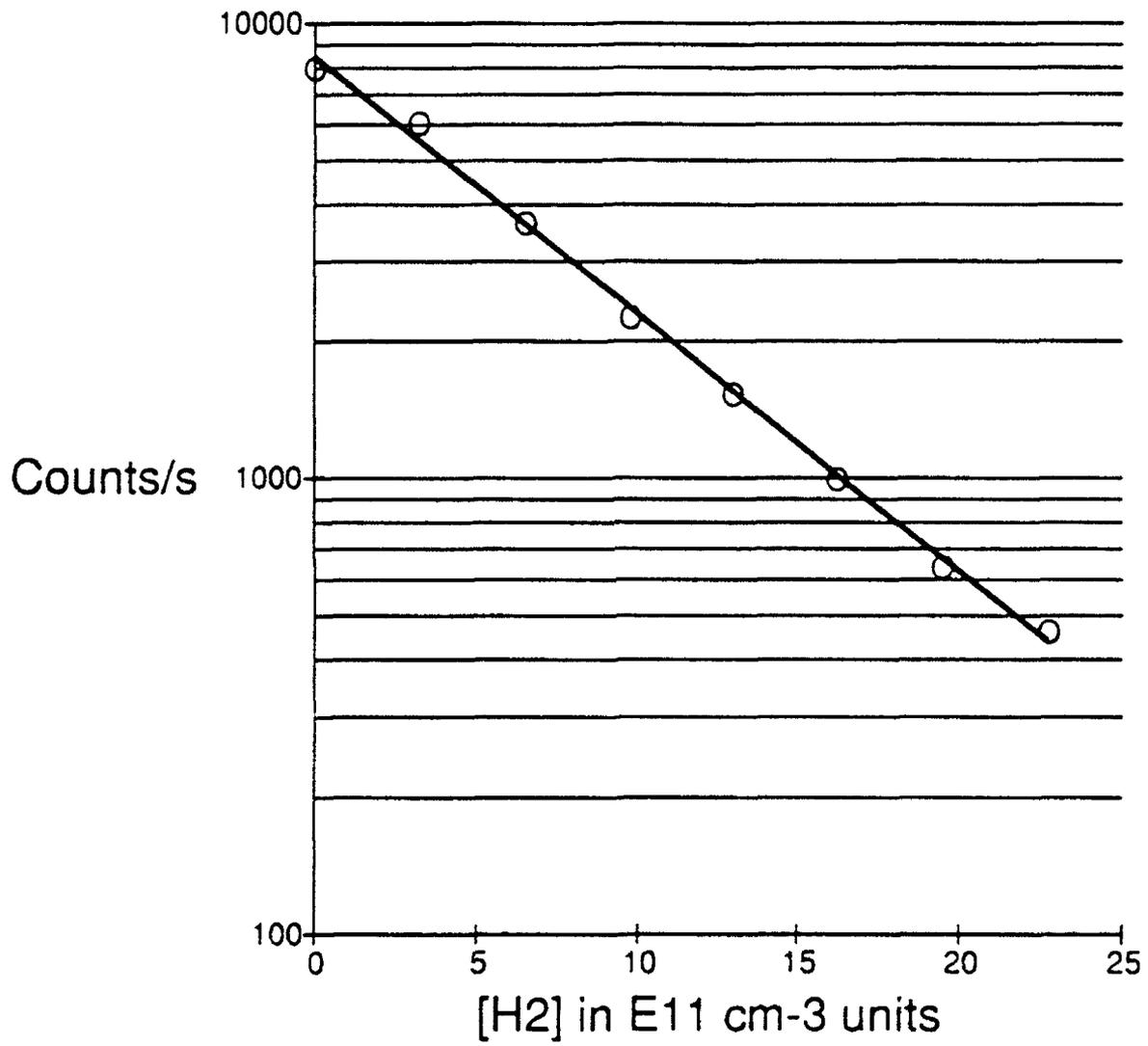


Fig. 2. Results for Cl⁻ + CH₃Br.

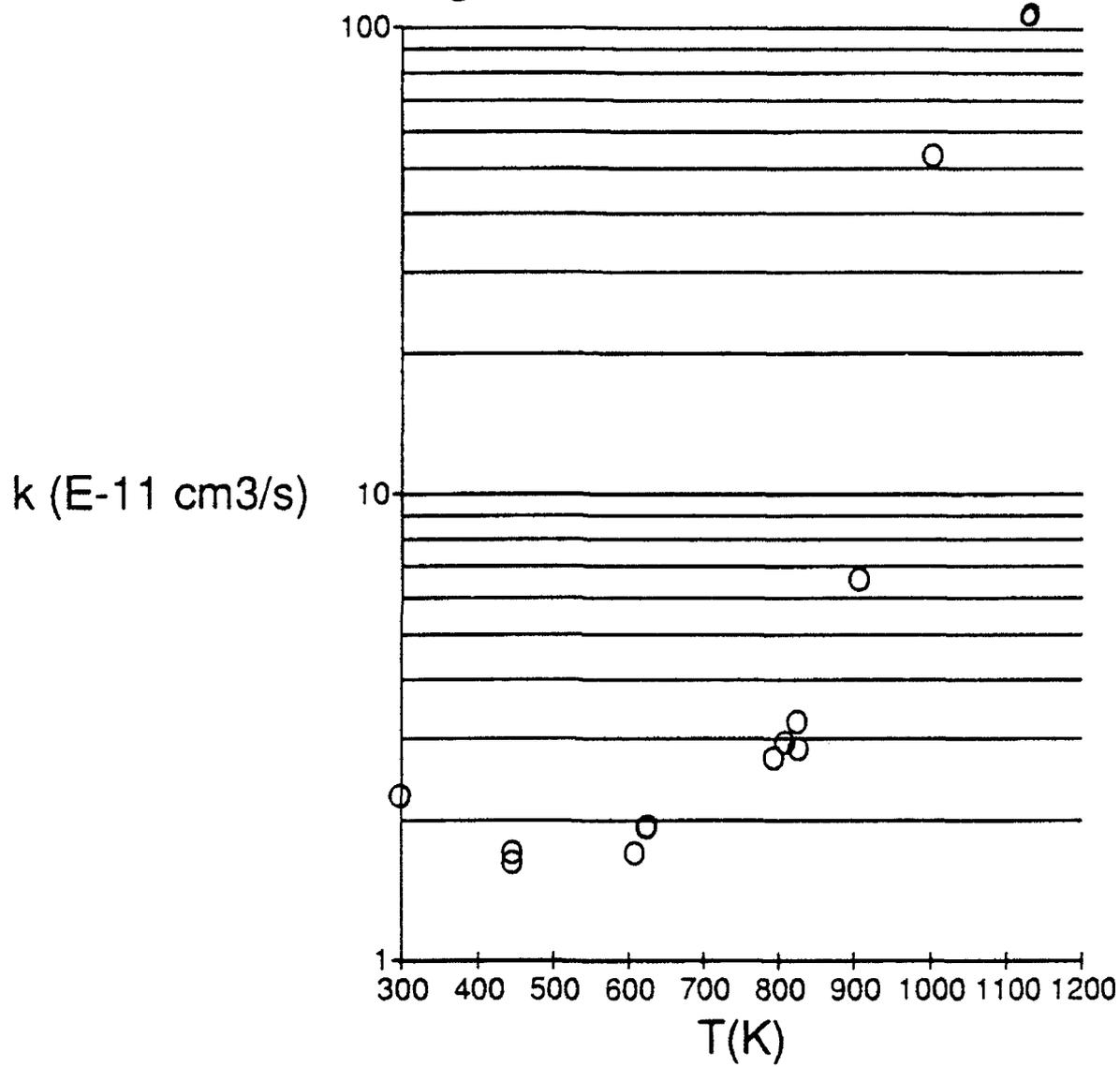


Fig. 3. Results for O- + CO, H₂, D₂, NO, and CH₄.

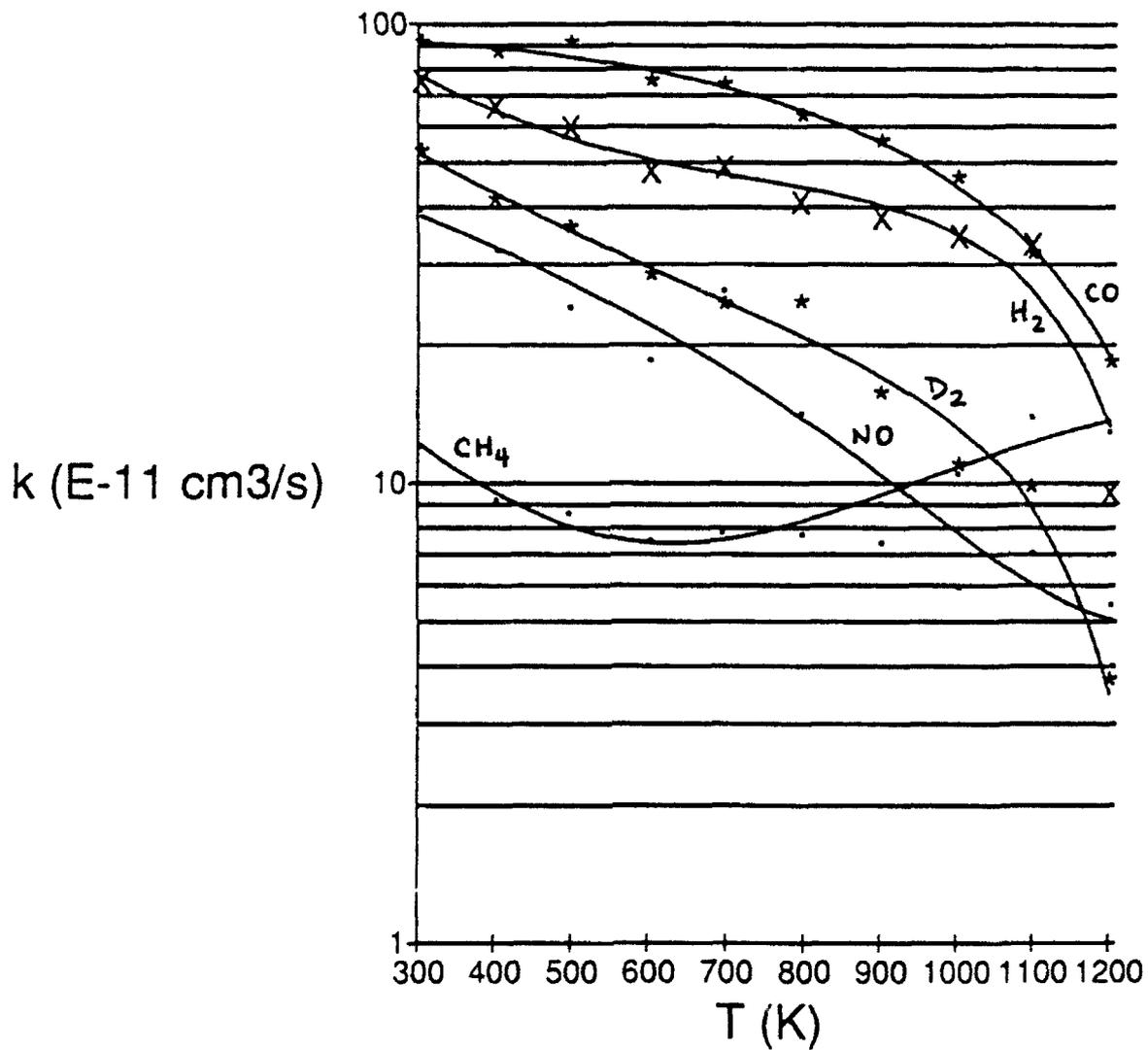
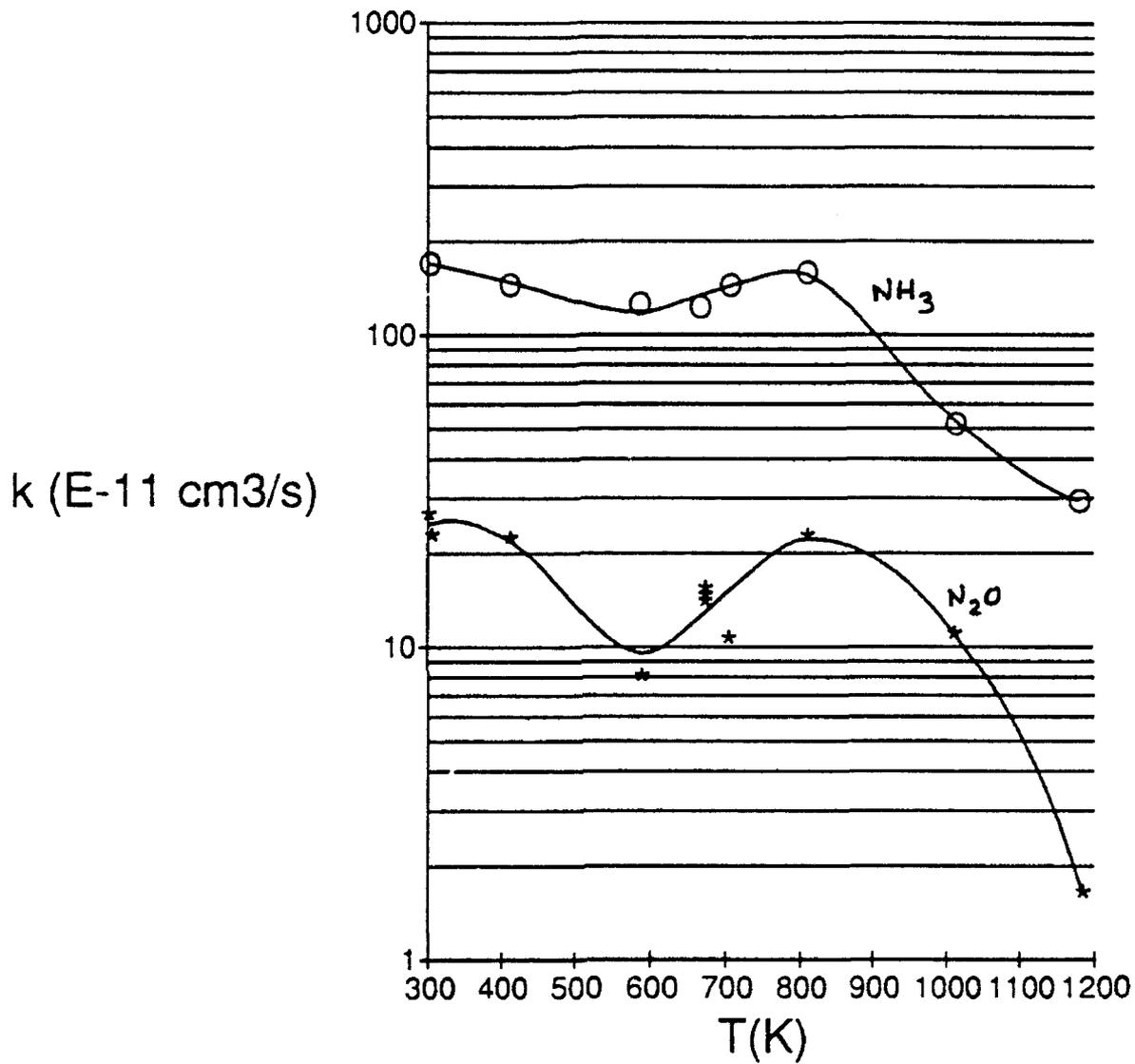


Fig. 4. Results for O- + NH₃, N₂O.



PARTICLE SIMULATIONS OF PLASMAS

R. D. Murphy
Professor
Department of Physics

University of Missouri - Kansas City
Kansas City, Missouri 64110

Final Report for:
Summer Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

September 1992

PARTICLE SIMULATIONS OF PLASMAS

R. D. Murphy
Professor
Department of Physics
University of Missouri - Kansas City

Abstract

Studies of the applicability of "particle" computer simulations to plasma phenomena were performed. A number of problems for which such simulations appear to be feasible were identified. A direct simulation Monte Carlo code, necessary for such efforts, was written and tested.

PARTICLE SIMULATIONS OF PLASMAS

R. D. Murphy

INTRODUCTION

The High Energy Plasma Physics Division of the Phillips Laboratory is a leading center in the investigation of the physics of dense, hot plasmas, an area of considerable importance to the Air Force. A major reason for this preeminent position is the Laboratory's SHIVA Star 9.5 Megajoule fast capacitor bank, which has made possible a number of experimental studies of plasma phenomena. Of particular importance for the present report on the author's summer activities are two current investigations: Magnetically Accelerated Rings for the Achievement of Ultra-high Density, Energy and Radiation (MARAUDER) and a project to achieve ultrahigh compression using electromagnetically imploded aluminum liners.

Another important reason for the Laboratory's preeminence in plasma physics investigations is the existence of a Computational Plasma Physics group and the accessibility to this group of state-of-the-art computer facilities at the Air Force Supercomputer Center as well as access to massively parallel machines. The collaboration and close interaction between theory and experiment has greatly benefitted both.

DISCUSSION AND METHODOLOGY

Plasma physics is characterized by a considerable variety of phenomena which operate at very different densities and on very different time scales¹ which range from sub-picosecond to greater than microsecond

times. This diversity has necessitated a wide variety of computational approaches, which can be roughly divided into "continuum" approaches, e.g. magnetohydrodynamics (MHD), and "particle" approaches, e.g. particle-in-cell (PIC), molecular dynamics (MD) and Monte Carlo methods. For a variety of reasons, especially the existence at Phillips Laboratory of state-of-the-art MHD codes and the author's own experience in implementing particle simulations, it was immediately evident that the most fruitful approach would be to identify problems of interest to the Laboratory which could be done best (or only) with particle codes, to study the feasibility of implementing particle codes and to proceed in such implementation as far as the time allotted for the author's summer faculty appointment permitted.

Three areas, each with a number of interesting problems, were identified. They are of considerable complexity in that each is cross-disciplinary and will require study of, in addition to plasma phenomena, such areas as intermolecular potentials and atomic, molecular, solid-state and liquid-state physics. The areas are as follows:

1) Wall effects. The MARAUDER experiments produce an argon plasma toroid which accelerates past metal containing walls. Spectroscopic evidence² indicates that the argon plasma contains silicon, carbon and hydrogen contaminants. This suggests two interesting lines of investigation: a) the interaction of the plasma with the contaminants (a plasma and molecular physics problem) and b) the physics of the absorption, desorption, etc. of the contaminants under bombardment by the

plasma (a solid-state and surface-science problem).

2) Plasma-induced blackout during re-entry of space vehicles. This problem, of potential interest to Phillips Laboratory, involves rarefied neutral gas dynamics and plasma, condensed-matter, atomic, molecular and radiation physics.

3) Properties of liquid and solid aluminum: the interaction of the imploding aluminum liners with a) plasma and b) metal walls.

At least part of each of the above areas requires development of new computational approaches; in addition, it must be stressed that the implementation of particle approaches will be extremely computationally-intensive and is likely to require substantial computer resources. In essence, a hybrid code, which combines PIC or MD methods with methods like the Direct Simulation Monte Carlo Method (DSMC)³, is needed. The DSMC is a probabilistic procedure for solving the Boltzmann equation which was originally developed for the study of rarefied neutral gas flows; in most such applications, the collisions were treated as "hard-sphere" collisions. But the method has been generalized^{4,5} to include other kinds of events, e.g. ionization, molecular excitation, charge attachment, charge exchange, etc. Birdsall⁵ has emphasized that this generalization and the merger of these codes is a developing art which relies heavily on experimental cross-section data.

Results and Conclusions

The principal conclusion of the author's study is that there are a number of areas, as noted above, which are pertinent to the work at the Phillips Laboratory and which can be addressed by particle simulation methods. These problems are not, however, trivial: they are interdisciplinary, they are computationally intensive, and they will require additional code development.

The major computational accomplishment of the summer's effort is the development by the author of a DSMC code which can be used on a "stand-alone" basis to study problems in rarefied gas dynamics and which can also be integrated with MD codes to produce a hybrid code. The DSMC code was tested against a standard aerodynamic case, a two-dimensional hypersonic wake trailing a blunt body⁶. A further and more stringent test of the code, an argon shock-wave density profile, is presently in progress. Although additional tests can and will be done to verify and refine the code, it is believed to be working correctly.

REFERENCES

1. See, for example, T. Tajima, Computational Plasma Physics, Addison-Wesley Publishing Co., New York (1989).
2. T. Englert, D. W. Price and G. F. Kiuttu, "Temporal and Spatial Disbribution of Contaminant Plasma in Compact Toroid Formation at the Phillips Laboratory," Paper presented at ICOPS 92.
3. G. A. Bird, Molecular Gas Dynamics, Clarendon Press, Oxford (1976).
4. C. Borgnakke and P. S. Larsen, J. Computational Phys., 18, 405 (1975).
5. C. K. Birdsall, IEEE Trans. on Plasma Sci., 19, 65 (1991).
6. D. Rault and E. P. Szuszezewicz, SAIC Scientific Report #17, October 1988 (unpublished).

POINTING AND ACCURACY ANALYSIS
OF SOLAR CONCENTRATORS

Rupa Purasinghe
Associate Professor
Department of Civil Engineering

California State University at Los Angeles
5151 State University Drive
Los Angeles, California 90032 -8151

Final Report for
Summer Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Edwards Air Force Base, Edwards, California

Focal Point
Kristi Laug

September 1992

POINTING AND ACCURACY ANALYSIS OF SOLAR CONCENTRATORS

Rupa Purasinghe
Associate Professor
Department of Civil Engineering
California State University at Los Angeles

Abstract

Under the space environment the paraboloid solar concentrators and support structures can deform and hence the focal point of the concentrators can diffuse. If this diffusion is large, energy will not concentrate on the thruster as desired. This report addresses this aspect of pointing and accuracy analysis of solar concentrators.

A simplified finite element model including support struts and simple torus were used. The torus model was made up of several equal length beams. The simple model did not contain the paraboloid reflector, and assumes the reflector does not affect the deformation of torus. Modal and buckling analysis of the structure and a static analysis due to equivalent thrust loads were done using the finite element Nastran program. The results demonstrate diffusion of the focal point of the concentrator.

POINTING AND ACCURACY ANALYSIS OF SOLAR CONCENTRATORS

Rupa Purasinghe

INTRODUCTION:

The Solar Propulsion Concept (SPC), consists of concentrators, solar energy absorbing thruster, and a single fuel tank. The SPC has the advantage of doubling specific impulse and hence doubling payload. This system can make Low Earth Orbit (LEO) to Geosynchronous Earth Orbit (GEO) missions less expensive. A light weight Paraboloid concentrators and a support structure is needed for the Solar Propulsion Concept [1]. Studies have shown that inflatable paraboloids offer great savings in payload and packaged volume compared to paraboloids mechanically erected in space [2,3]. The reflectors and support structure should be packageable within the launch vehicle and must be deployable once Low Earth Orbit is achieved. Leakage through holes caused by meteoroids is easily compensated by make up gas, due to a very low pressure requirement. Other advantages of the inflatable models are improved dynamic performance through rapid damping, non linear restoring forces, and reduction of thermal gradients [4].

Work on inflatable reflectors has been done for over 25 years. Paraboloids were built and tested for surface accuracy and microwave performance [4]. For solar concentrators, slope error is the key accuracy parameter for measuring gross surface distortion in construction. This has been addressed by M. Thomas [5].

Another aspect of accuracy of solar concentrators under the space environment is the concentrator deflection and rotation due to thrust and temperature. Deflection and rotation tends to make the focal point too diffuse. If the diffusion of the focal point is large, energy will not concentrate on the thruster as desired. This report addresses this aspect of pointing and accuracy of focal point diffusion in concentrators.

For this study two off axis paraboloid reflectors supported on a rigidized Elliptical Torus were selected. Each torus was supported on three struts which

are fixed to a turntable [Figure 1]. It is assumed that the paraboloid reflectors were attached to torus in such a way that reflectors do not affect the torus. This system was analyzed for equivalent thrust loads in space.

METHODOLOGY:

A finite element analysis of the Concentrator support structure was done using the NASTRAN COSMIC Program. The finite element model of the support structure is shown in Figure 2. It consists of beam elements (CBAR) for both struts and torus. The torus was modelled with 24 straight beam elements. The struts are fixed to the turntable. The weight of the mirror is lumped at torus nodal points.

VIBRATIONAL ANALYSIS:

A modal analysis of the structure was done using the Finite Element Analysis with struts being fixed to turntable as described above. The first few frequencies for the structure is shown below. As noted below the frequencies are small. Figures 3 to 5 shows first few mode shapes. It should be noted that the three struts are modeled with three elements each (only for modal analysis), and hence only two internal nodal point displacements are possible.

MODE	FREQUENCY (Hz)
1	0.0036
2	0.0696
3	0.1362
4	0.2012
5	0.3492
6	0.3721

LINEAR BUCKLING ANALYSIS:

A linear buckling analysis was done using the model in Figure 2. The critical buckling load was calculated as a multiple of equivalent static load in X and Y directions. The buckling analysis is done using two subcases. One

with static 0.002g gravity in given direction, and the other with buckling analysis with differential stiffness[6]. The buckling load is the eigen value of analysis multiplied by static load value applied for sub case 1 [6].

$$\begin{aligned}\text{The Buckling Load due to load in X direction} &= \text{Eigen Value} * \text{Load} \\ &= 19.982 * 0.002 \text{ g} \\ &= 0.03996 \text{ g}\end{aligned}$$

$$\begin{aligned}\text{The Buckling Load due to load in Y direction} &= \text{Eigen Value} * \text{Load} \\ &= 7.398 * 0.002 \text{ g} \\ &= 0.0146 \text{ g}\end{aligned}$$

This shows that the possibility of buckling is remote due to equivalent thrust load of 0.002 g in X or Y directions.

FINITE ELEMENT MODEL:

The basic finite element model is shown in Figure 2. In order to find the focal point displacement with reference to displacement of torus, rigid elements were connected between points on torus and dummy nodes at focal points. The rigid elements were CRIGID2 and CRIGID elements. These elements have no stiffness values.

The displacement of focal point is calculated as follows:

- i) Displacement due to displacement on torus.
- ii) Twice displacement due to rotation of torus.

This is because the torus/reflector will cause light rays to be reflected at twice the angle of torus rotation.

The CRIGID2 elements provides displacements at focal point due to displacement and rotation of points on torus. On the other hand, CRIGID elements provides displacements due to displacements alone at points on torus. The results of this analysis is shown in Table 1 and 2.

Nodal Coordinates of Three Struts in inches

Strut 1	(0.0, -12.0, 18) to (0.0 -520.8, 514.2)
Strut 2	(10.39, 6.0, 18.0) to (540.96, 208.9, 1341.7)
Strut 3	(-10.39, 6.0, 18.0) to (-540.96, 208.9, 1341.7)

Center of Torus (0.0,0.0,1104.8)

Focal Point of Paraboloid Mirror (0.0,0.0,12.0)

In modelling the frame, X,Y,Z coordinate system was used for three struts, and cylindrical coordinate system for 24 nodal points on torus with equal length straight CBAR elements.

Materials Used:

Space Cured Gr/Ep composite with M40J fibers

+ - 30 / 0 / 90 / 30 unidirectional tape

Density = 0.059 lb/in³ density

Modulus of Elasticity = 17,210 ksi

Diameter of Torus = 13 in (0.015 in thickness)

Diameter of Struts = 7 in (0.015 in thickness)

Moment of Inertia of three struts = 2.02 in⁴

of torus = 12.8 in⁴

Polar Moment of Inertia of struts = 4.04 in⁴

of torus = 25.6 in⁴

TABLE 1

Focal Point Displacements with reference to Fixed Turntable

(inches)

(Due to 0.002 g in X direction)

X direction	Y direction	Z direction
-7.575	2.022	4.094
-6.942	14.632	3.099
-2.528	5.343	0.985
-1.793	3.011	0.591
-2.007	1.860	0.447
-2.573	0.949	0.282
-5.669	0.000	-0.000
-2.573	-0.949	-0.282
-2.007	-1.860	-0.447
-1.793	-3.011	-0.584
-2.528	-5.342	-0.985
-6.943	-14.632	-3.099
-7.575	-2.022	-4.093
-6.484	15.832	-4.099
-5.063	7.806	-3.480
-4.091	5.015	-2.641
-3.762	3.219	-1.870
-3.862	1.550	-1.118
-7.881	0.000	0.000
-3.827	-1.550	1.118
-3.762	-3.219	1.870
-4.091	-5.015	2.641
-5.063	-7.806	3.480
-6.484	-15.832	4.099

TABLE 2

Focal Point Displacement with reference to Fixed Turntable
(inches)

(Due to 0.002 g in Y direction)

X direction	Y Direction	Z Direction
-0.600	-0.374	0.324
-0.139	-1.012	0.125
1.605	-1.192	0.421
3.972	-1.029	-0.802
5.475	-0.434	-0.649
7.515	0.181	-0.306
-0.000	0.443	-0.137
-7.515	0.181	-0.306
-5.475	-0.434	-0.649
-3.972	-1.029	-0.802
-1.614	-1.192	-0.417
0.139	-1.011	0.125
0.601	-0.374	0.324
0.511	-0.282	0.262
0.288	-0.430	0.619
0.131	-0.500	-0.177
0.091	-0.514	-0.360
0.142	-0.484	-0.455
0.000	-0.455	-0.472
-0.142	-0.483	-0.455
-0.091	-0.514	-0.360
-0.131	-0.500	-0.177
-0.281	-0.430	0.062
-0.511	-0.282	0.262

The above 24 displacement values are calculated with reference to displacement/rotation of nodal point of torus.

reflectors stiffen the torus, there by controlling displacements and rotations ?

To address this issue, the inflated reflectors and support structure, as a whole, need to be analyzed using finite elements as shown in Figure 6. The Paraboloid reflector can be modeled by PATRAN[8] preprocessor, and the analysis done with NASTRAN Finite Element program. This will allow better understanding of the Pointing and Accuracy issue of Solar concentrators.

DISCUSSION:

This report documents the possibility of diffusion of focal point on solar concentrators due to thrust loads. The values documented suggest non uniform diffusion of focal point with relation to points on torus. If stiffer members are selected for torus, it is possible to control deflections of focal point. A complete analysis of the reflector and strut system will allow more accurate analysis.

ACKNOWLEDGEMENT:

Financial support through the Research and Development Laboratory of AFOSR is appreciated. Thanks are due to the focal point Kristi Laug, and M. Holmes, S. Cartier and Capt. D. Castillo for their help during this study.

REFERENCES:

- 1) K. K. Laug, " The Solar Propulsion Concept is Alive and Well at the Astronautics Laboratory," Joint Army, Navy, NASA, Air Force Propulsion Meeting, 1989, Cleveland, Ohio.
- 2) M. Thomas and G. J. Friese, " Pressurized Antennas for Space Radars," AIAA Collection CP 807, from the AIAA Sensor Systems for the 80's Conference, December 1980.
- 3) G. J. Friese, G. D. Bilyeu, and M. Thomas, " Initial 80's Development of Inflated Antennas," L'Garde Report LTR-82 -GF-107, NASA Contractors Report, December 1982.
- 4) A. J. Wendt and L. D. Surber, " Inflatable Antennas", 3rd Aerospace

Expandable and Modular Structures Conference, AFAPL TR 68-17, Mar 67.

5) M. Thomas and G. Veal, " Scaling Characteristics of Inflatable Paraboloid Concentrators", Solar engineering, ASME 1991.

6) NASTRAN MANUAL, NASTRAN Finite Element Software.

7) Lockheed Missiles and Space Company - Status Report, July 92.

8) PATRAN Software- PDA Engineering

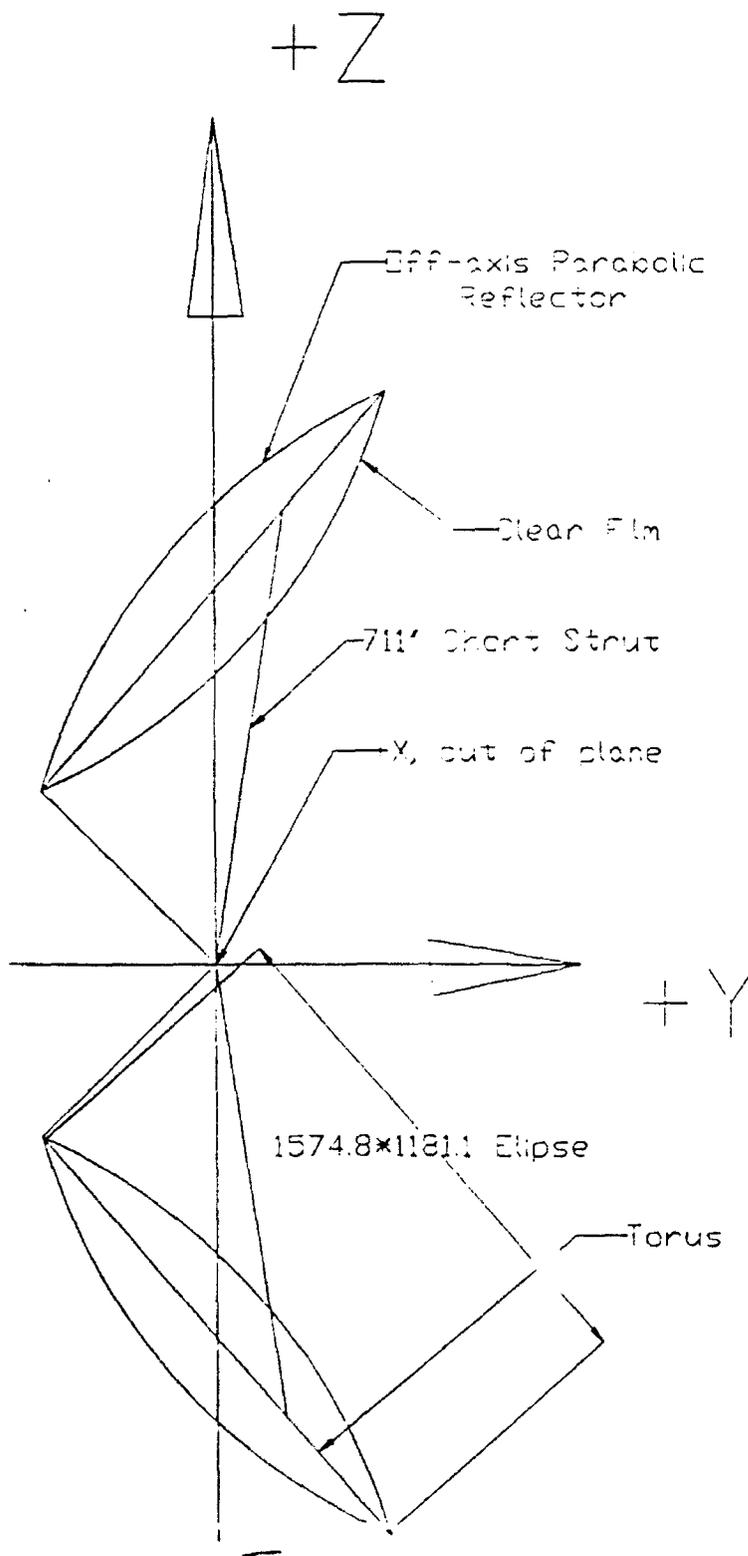


Figure 1: Solar Concentrator System

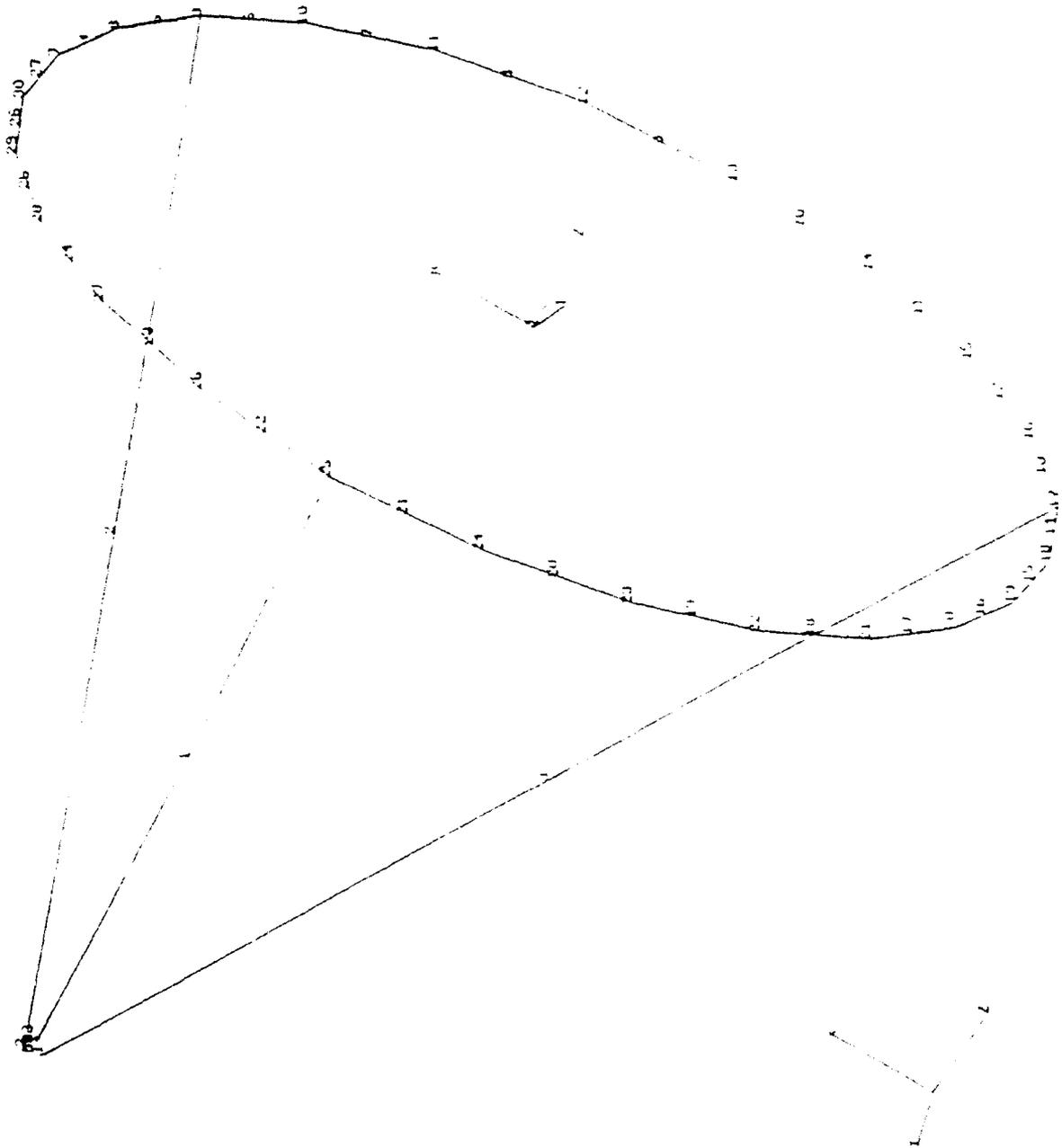


Figure 2: Nastran Finite Element Model
26-13

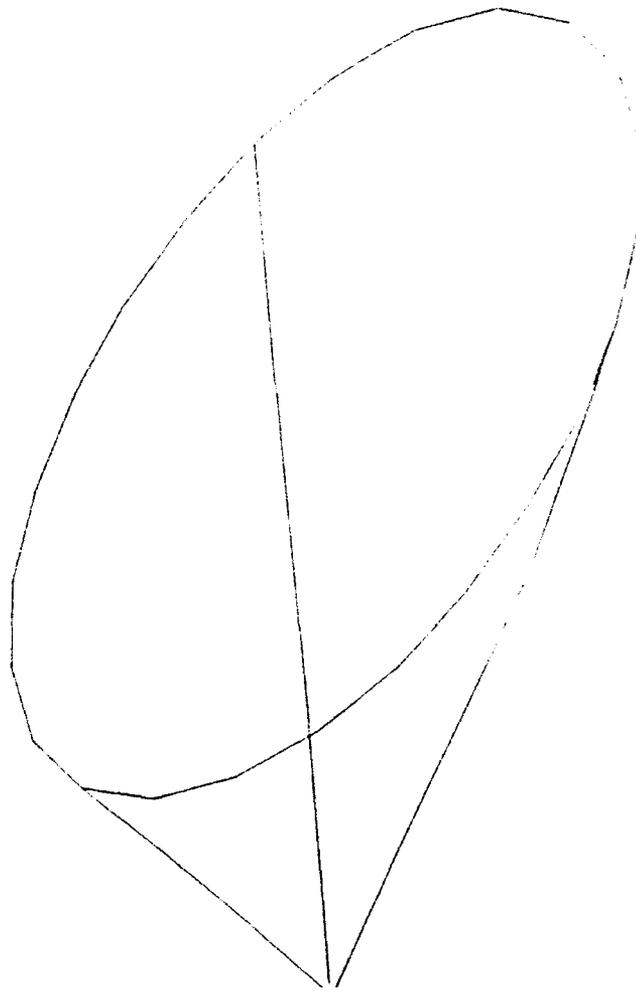


Figure 3: First Mode of Dynamic Analysis

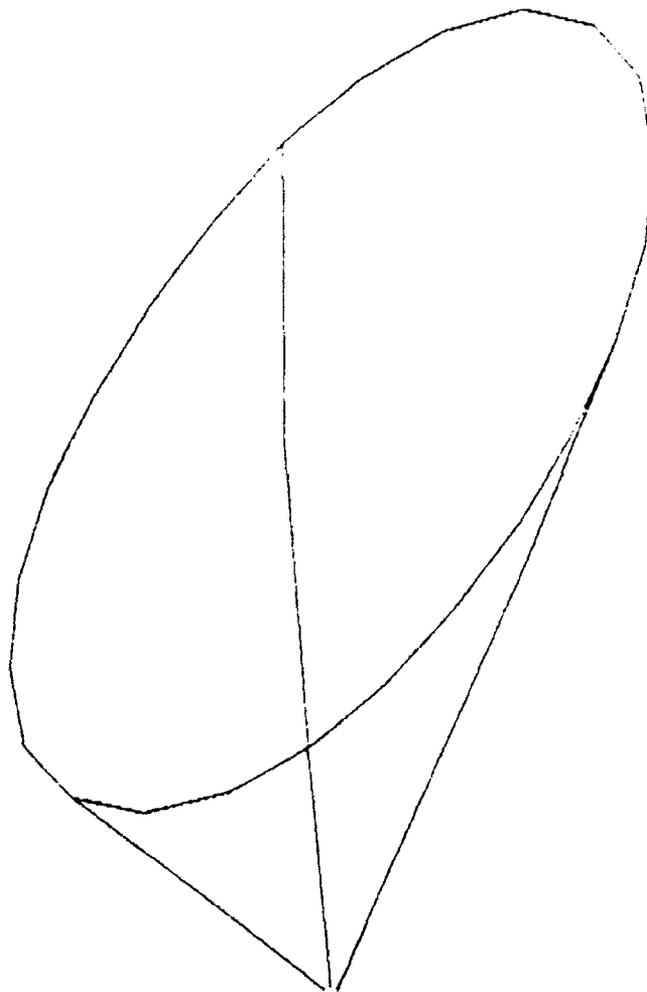


Figure 4: Second Mode of Dynamic Analysis
26-15

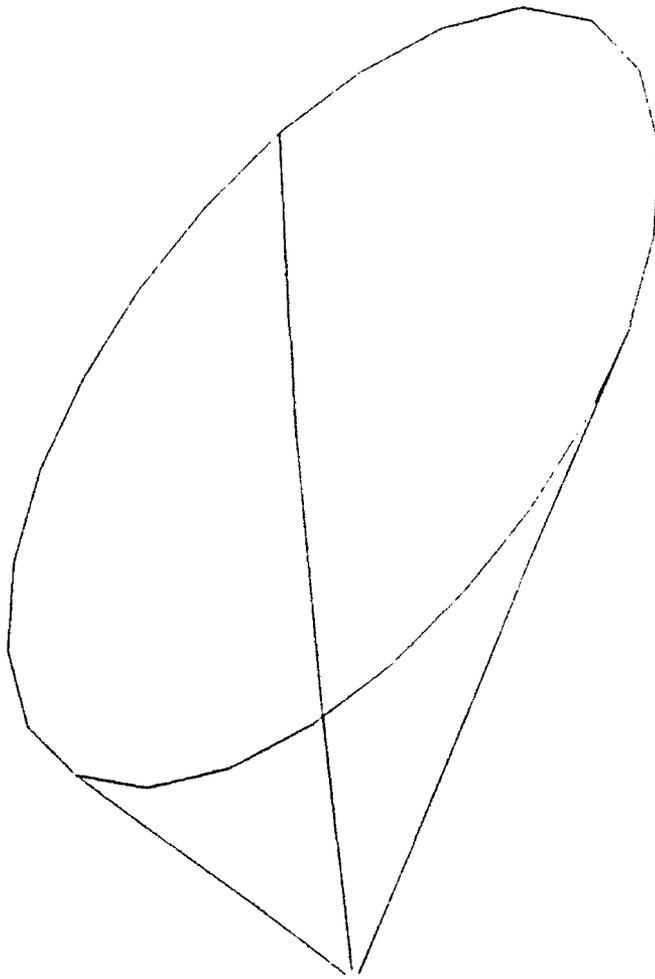


Figure 5: Third Mode of Dynamic Analysis

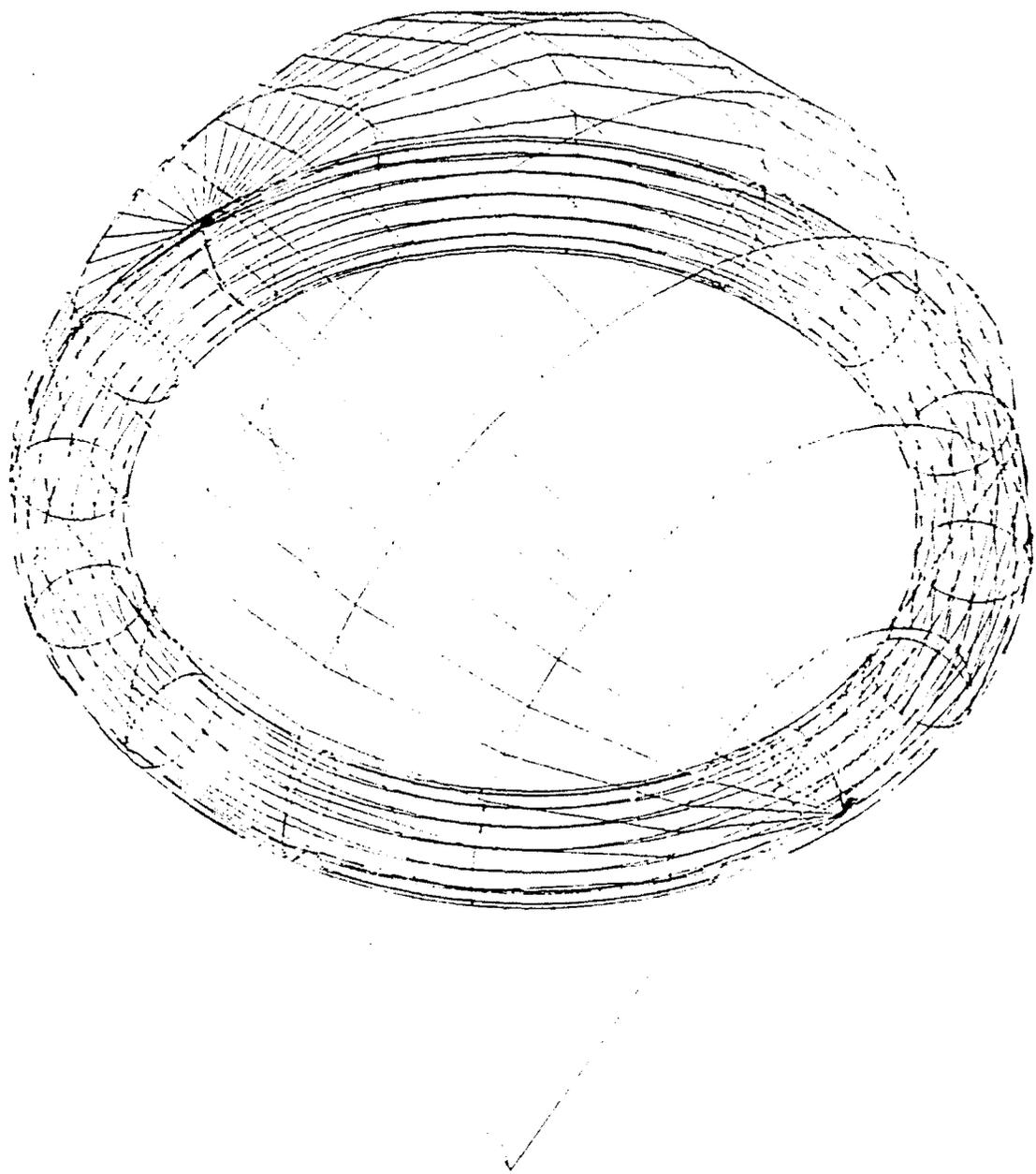


Figure 6: Proposed Analysis of Concentrator System
(not to scale)

MIXED-MODE FRACTURE OF SOLID PROPELLANTS

**Krishnaswamy Ravi-Chandar
Associate Professor
Department of Mechanical Engineering**

**University of Houston
Houston, TX 77204-4792**

**Final Report for:
Summer Research Program
Phillips Laboratory**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington DC**

September 1992

MIXED-MODE FRACTURE OF SOLID PROPELLANTS

Krishnaswamy Ravi-Chandar
Associate Professor
Department of Mechanical Engineering
University of Houston
Houston, TX 77204-4792

Abstract

The initiation of crack growth under mixed-mode loading conditions in a solid propellant material was examined experimentally. Compact tension-shear specimens were prepared from a solid propellant. The specimens were loaded under combined tension and shear in a special grip capable of varying the ratio of tension and shear. The crack tip deformation was examined using a coarse grid (0.2 mm spacing) painted on the surface of the specimen and photographed with a video camera before and during the loading. While the critical stress intensity factors at crack initiation were found to correlate well with the maximum tangential stress theory, the angle of crack kinking did not. This is attributed to extensive crack blunting and damage near the crack tip region.

MIXED-MODE FRACTURE OF SOLID PROPELLANTS

Krishnaswamy Ravi-Chandar

1 INTRODUCTION

Solid propellant rocket motors are subjected to a various loading conditions during manufacture, storage, transport and firing. Cracks appear in rocket propellants at locations of stress concentration due to these loads as shown in Figure 1. In order to determine the reliability of cracked solid propellant materials, it is necessary to investigate the fracture behavior of these materials. A substantial amount of research has been performed in this area over the last thirty years. A survey of these quickly indicates that the most common assumption involved in the investigations is that the loading is symmetrical with respect to the crack line and thus the crack is under a mode-I loading condition; however, under arbitrary loads and crack geometry, it is possible that the crack will be under unsymmetrical loading, combining both mode-I and mode-II loads. In this paper, we explore the problem of fracture under mixed-mode loading in solid propellants. In Section 2, the basic linear elastic fracture mechanics characterization of mixed mode fracture is reviewed along with the different criteria that have been proposed in the literature. The experimental scheme used in the present investigation is discussed in Section 3, followed by an examination of the results in Section 4. The results of this study indicate that the problem of mixed-mode fracture in solid propellants is complicated by the fact that large deformations and damage occur at the crack tip region and point to the need for further investigation.

2 THE MIXED-MODE FRACTURE PROBLEM

The problem of mixed-mode fracture was initially addressed by Erdogan and Sih [2] in conjunction with the questions of whether the second and third modes of fracture really existed and whether critical stress intensity factors for pure modes II and III had any physical meaning. In the two-dimensional case, the Cartesian components of the crack tip stresses can be expressed (in terms of the polar coordinates as)

$$\sigma_{\alpha\beta} = \frac{K_I}{\sqrt{2\pi r}} f_{\alpha\beta}(\theta) + \frac{K_{II}}{\sqrt{2\pi r}} g_{\alpha\beta}(\theta) \quad (1)^1$$

where $f_{\alpha\beta}(\theta)$ and $g_{\alpha\beta}(\theta)$ are characteristic functions and K_I and K_{II} are the mode-I and mode-II stress intensity factors respectively. The asymmetry of the loading is generally described in terms of the mixity parameter μ which is defined as

$$\mu = \frac{K_{II}}{K_I} \quad (2)$$

Thus the mixed-mode fracture problem involves determining critical stress intensity factor pairs (K_I^c, K_{II}^c) for different values of μ , at which the crack will initiate. Furthermore it is required to determine the crack initiation angle γ as a function of μ as well as the subsequent crack propagation path. The different criteria that have been proposed to

¹ Index notation is used; α, β have the range 1,2.

predict the critical condition for crack initiation and the crack trajectory upon initiation are reviewed in the following paragraphs.

2.1 The Maximum Tangential Stress Criterion (MTS)

This criterion was proposed by Erdogan and Sih [2] and it is stated as follows: *Crack extension starts at the crack tip in radial direction. This extension is in that radial direction perpendicular to the direction of the greatest tension. Crack extension begins when this tension reaches a certain critical value at a certain distance from the crack tip.* Erdogan and Sih used the center-cracked panel configuration to simulate the mixed-mode problem with a crack of length $2a$ subjected to uniaxial far-field loading σ , the crack being oriented at angle β to the loading direction. They obtained a relation for the initiation angle γ and the loading mixity, $\mu = \tan \beta$, by equating the partial derivative $(\partial\sigma_{\theta\theta}/\partial\theta)$ to zero (equivalently the shear stress $\sigma_{r\theta}$ can be equated to zero). Both these conditions yield the relation

$$\sin \gamma + \mu(3 \cos \gamma - 1) = 0 \quad (3)$$

From (3) it can be shown that for pure mode-II conditions ($\mu = 0$) the initiation angle is predicted to be 70.5° . A limited amount of experimental evidence that showed good agreement with the predictions of the MTS criterion was presented by Erdogan and Sih. Williams and Ewing [3] conducted a more detailed experimental investigation of the same problem using PMMA plates and confirmed the previous results in the same range of β used by the previous authors ($30^\circ \leq \beta \leq 80^\circ$). However they found that in the range of $\beta < 15^\circ$ there was a marked deviation from the predictions of the MTS theory and the initiation angle tended more towards 90° than towards 70.5° as one approached pure mode-II conditions. They were also able to show that a better correlation with experiment could be obtained if the crack parallel stress component was introduced in the analysis. In including this crack parallel component Williams and Ewing also introduced a critical distance (r_c), from the crack tip, at which the $\sigma_{\theta\theta}$ stress was postulated to be a maximum. In a further discussion of the same work, Finnie and Saith [4] detected an error in the analysis of Williams and Ewing, which they corrected and were able to show a better correlation of experimental data with the predictions of the corrected analysis.

2.2 The Strain Energy Density Criterion (SED)

This theory was proposed by Sih [5] and it involves the following two hypotheses: *The crack extends in a radial direction that is characterised by the minimum strain energy density. Crack extension begins when the strain energy density in this particular direction reaches a critical value.* The strain energy density is defined in terms of K_I and K_{II} as

$$S = a_{11}K_I^2 + a_{12}K_I K_{II} + a_{22}K_{II}^2 \quad (4)$$

where a_{11} , a_{12} and a_{22} are functions of θ and are defined according to the conditions of plane stress and plane strain, in terms of the Poisson's ratio.

A limited amount of experimental data was presented in support of this criterion in the original paper. A feature

of this criterion which makes it different from the others, is that it incorporates the influence of a material parameter in the form of the Poisson's ratio. In a study of this criterion, in comparison to the MTS criterion, Finnie and Weiss [6] conducted experiments on cross-rolled beryllium sheets that exhibited an effective Poisson's ratio equal to zero. The results from these experiments were more in agreement with the MTS criterion than with the SED criterion.

Apart from the above mentioned, four more criteria viz the maximum potential energy release rate criterion, the Maximum Tangential Strain criterion (MTSN) based on the near tip strains, the Energy Momentum Tensor criterion based on Eshelby's energy momentum vector, and the T-criterion have been proposed by Palaniswamy and Knauss[7], Wu [8], Tirosh [9] and Theocaris [10] respectively. However, in accordance with the trend shown in literature, attention is focussed on the first two criteria.

3. EXPERIMENTAL METHOD

The compact-tension-shear specimen configuration was chosen as the most suitable for mixed-mode fracture investigations of solid propellants; the major reasons for this choice are discussed fully by Mahajan and Ravi-Chandar [11]. Figure 2 shows a diagram of the experimental scheme; the propellant was bonded to the aluminum end tabs using a urethane polymer and the end tabs were attached to the aluminum loading fixtures using bolts. By proper selection of diametrically opposed loading holes, this assembly provides a wide range of mode mix μ . The specimens were loaded in an Instron tensile testing machine until initiation of crack growth. The stress intensity factors K_I and K_{II} for this configuration can be determined from the numerical simulation of Buchholz et al [12], once the loading is monitored through the load cell on the testing machine. In addition to this method of determining the crack tip conditions, two other methods were investigated; first, a fine square grid with 0.2mm spacing was deposited on the specimen. This grid was photographed using a video camera both before and during the loading of the specimen; from these pictures, the displacement and strain fields ahead of the crack tip can be determined. The results of this are discussed in the next section. The second method that was proposed to be used was the method of caustics [13]. In this method, a parallel beam of light is reflected from the surface of the specimen; due to the surface deformation near the crack tip, light rays from the crack tip region are deviated away from their straight path and form a caustic curve on a virtual screen at some distance behind the specimen. In order to implement this technique, it was necessary to coat the propellant specimen with a reflective coating. Since attempts to accomplish this were not successful, only the grid method of determining the crack tip strain field was utilized in the present investigation.

Nine propellant specimens were tested at different loading angles to investigate the initiation of crack growth under mixed mode loading. First, the deformation field near the crack was investigated using the results of the grid measurements; this was essential in determining the applicability of the linear elastic fracture mechanics characterization of the failure criterion using the concepts discussed in this section. These results are discussed in Section 4. Then in Section 5, the criterion for initiation of mixed mode cracking in solid propellants is discussed.

4. CRACK TIP DISPLACEMENT FIELD

The video pictures of the grids coated on the specimen were analyzed to determine the displacement field in the

vicinity of the crack tip. Examples of the deformed grid pattern are shown in Figure 3 just prior to crack initiation (Figure 3a) and after substantial amount of crack extension along a kinked path (Figure 3b). In order to determine the suitability of the LEFM formulation, a comparison of the displacement field given by

$$u_{\alpha} = K_I \sqrt{\frac{r}{2\pi}} \tilde{u}_{\alpha}^I(\theta) + K_{II} \sqrt{\frac{r}{2\pi}} \tilde{u}_{\alpha}^{II}(\theta) \quad (5)$$

and the measured displacement field must be performed. $\tilde{u}_{\alpha}^I(\theta)$ and $\tilde{u}_{\alpha}^{II}(\theta)$ are known functions of θ . The simplest comparison of the experimentally measured displacement and the theoretically computed displacement is to normalize the displacement component by \sqrt{r} and follow the angular variation. In other words, $\tilde{u}_{\alpha}^I(\theta)$ and $\tilde{u}_{\alpha}^{II}(\theta)$ are extracted from the experimental measurement and then compared with the analytical solution. For the case of pure mode-I, such a comparison is shown in Figure 4.

In Figure 4, $\tilde{u}_{\alpha}^I(\theta)$ is plotted as calculated from the experimental measurements at different radii from the crack tip. In determining the displacement components, a Lagrangian coordinate system was used and thus, $\tilde{u}_{\alpha}^I(\theta)$ was evaluated at a fixed distance r from the crack tip in the initial (undeformed) coordinate system. However, in the propellant specimens, the displacements are quite large and it is essential that the values of $\tilde{u}_{\alpha}^I(\theta)$ as evaluated at a fixed distance R in the deformed coordinate be used. This requires the development of an interpolating scheme; this is currently under development. In Figure 4, the comparison is made in the initial coordinate system and thus is appropriate only where the displacements are small; for the crack problem, this is typically in a small sector ahead of the crack and thus attention is focussed at an angular sector of $|\theta| \leq \pi/4$. From Figure 4, it can be observed that in this angular region, as the distance from the crack tip increases, the measured variation of $\tilde{u}_{\alpha}^I(\theta)$ agrees quite closely with the prediction from LEFM. Thus, one might argue that a linear elastic characterization of the failure behavior should be appropriate for this propellant material.

5. MIXED-MODE CRACK INITIATION CRITERION

From the nine experiments, the values of (K_I^f, K_{II}^f) at crack initiation were determined. These are displayed in Figure 5 as an interaction curve between $\frac{K_I^f}{K_{IC}^f}$ vs $\frac{K_{II}^f}{K_{IC}^f}$, where K_{IC}^f is the pure mode-I fracture toughness. Also shown in Figure 5 is the prediction of the maximum tangential stress criterion. As can be seen, there exists a large scatter in the experimental measurements. The scatter is attributable to variations in the material property due to inhomogeneities; since the solid propellant is a particulate composite with fairly large sized particles, small local variations in the spatial distribution of particles could lead to large variations in the measured toughness data. Wang and Beckwith [14] also report data on mixed-mode fracture with large data scatter on a propellant with a smaller percent content of solid particles. Their results indicate good agreement with the maximum tangential stress criterion, but if the strain energy density criterion is applied, the agreement is within the data scatter. Thus, to distinguish between the two criteria, it is necessary to consider the angle of crack kinking at initiation.

In Figure 6, the angle of crack initiation, γ , is plotted as a function of the mixed mode parameter $\delta = \mu/(1+\mu)$. Also shown in this figure are the predictions from the maximum tangential stress criterion and the strain energy density criterion. Clearly, neither of these two criteria are able to predict the crack kinking angle. Surprisingly, if the

angle of the line normal to the loading line in Figure 2 is considered, the crack kinking angle seems to be normal to this line. The implication of this result is that the crack kinking is not particularly influenced by the crack tip stress field. In other materials, the crack path initially kinks along the direction predicted by eq.(3) and then follows a curved path until a direction normal to the load line is reached [2,11]. As can be seen from Figure 3, the kinked crack path is straight. The large scale crack blunting and material damage in the crack tip region are expected to play a crucial role in determining the crack kinking angle and this needs to be investigated further.

6. CONCLUSIONS

The initiation of crack growth under mixed-mode loading conditions in a solid propellant material was examined experimentally. Compact tension-shear specimens were prepared from the solid propellant. The specimens were loaded under combined tension and shear in a special grip capable of varying the ratio of tension and shear. The crack tip deformation was examined using a coarse grid (0.2 mm spacing) painted on the surface of the specimen and photographed with a video camera before and during the loading. From the grid measurements, it was concluded that a square-root singular stress field was a suitable description of the crack tip deformation field. While the critical stress intensity factors at crack initiation were found to correlate well with the maximum tangential stress theory, the angle of crack kinking did not. This is speculated to be due to extensive crack blunting and damage near the crack tip region and needs to be investigated further.

REFERENCES

- [1] R. Sutton, Elements of Rocket Propulsion.
- [2] F. Erdogan and G.C. Sih, *Journal of Basic Engineering*, **85D**, (1963), p.519.
- [3] J.G. Williams and P.D. Ewing, *International Journal of Fracture*, **8**, (1972), p.441.
- [4] I. Finnie and A. Saith, *International Journal of Fracture*, **9**, (1973), p.484.
- [5] G.C. Sih, *Engineering Fracture Mechanics*, **5**, (1973), p.365.
- [6] I. Finnie and H.D. Weiss, *International Journal of Fracture*, **10**, (1974), p.136.
- [7] K. Palaniswamy and W.G. Knauss, *Mechanics Today*, Vol., **4**, (1978), p.87.
- [8] C.H. Wu, *Journal of Applied Mechanics*, **45**, (1978), p.553.
- [9] J. Tirosh, *Engineering Fracture Mechanics*, **9**, (1977), p.607.
- [10] P.S. Theocaris, G. Kardomateas and N.P. Andrianopoulous, *Engineering Fracture Mechanics*, **17**, (1982), p.439.
- [11] R.V. Mahajan and K. Ravi-Chandar, *International Journal of Fracture*, **41**, (1989), pp.235-252.
- [12] F.G. Buchholz, P.J.M Pirro, H.A. Richard and K.H. Dreyer, *Proceedings, Fourth International Conference on Numerical Methods in Fracture Mechanics*, (eds Luxmore *et.al*), Pineridge Press, Swansea, UK (1987), p.641.
- [13] J.F. Kalthoff, in *Handbook of Experimental Mechanics*, (ed. A.S. Kobayashi), Prentice Hall, 1989.
- [14] D.T. Wang and S.W. Beckwith, in 16th Meeting of the Mechanical Behavior Subcommittee of JANNAF, 1980, pp.241-254.

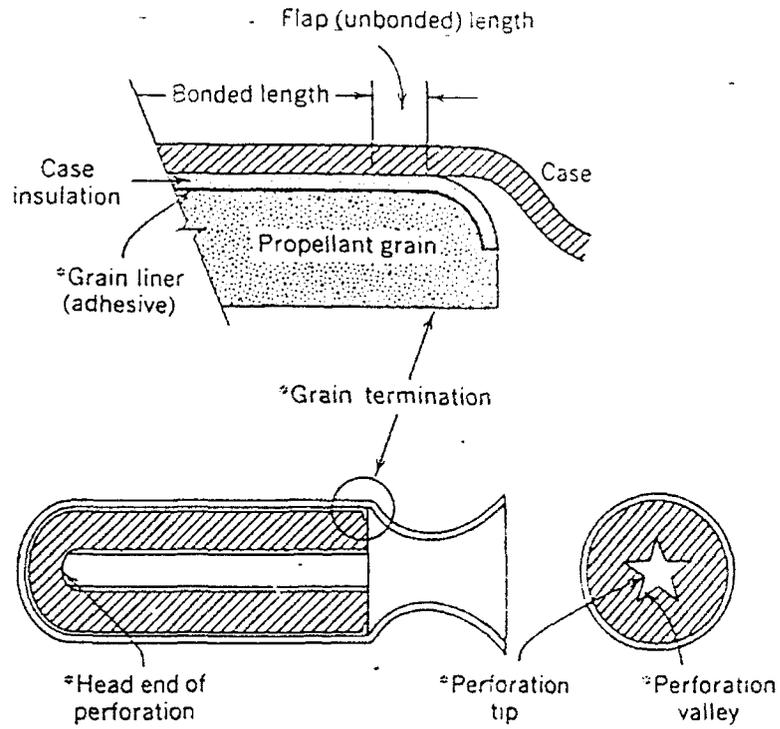


Figure 1. Detail of a solid rocket motor indicating points of potential cracking (from Sutton [1]).

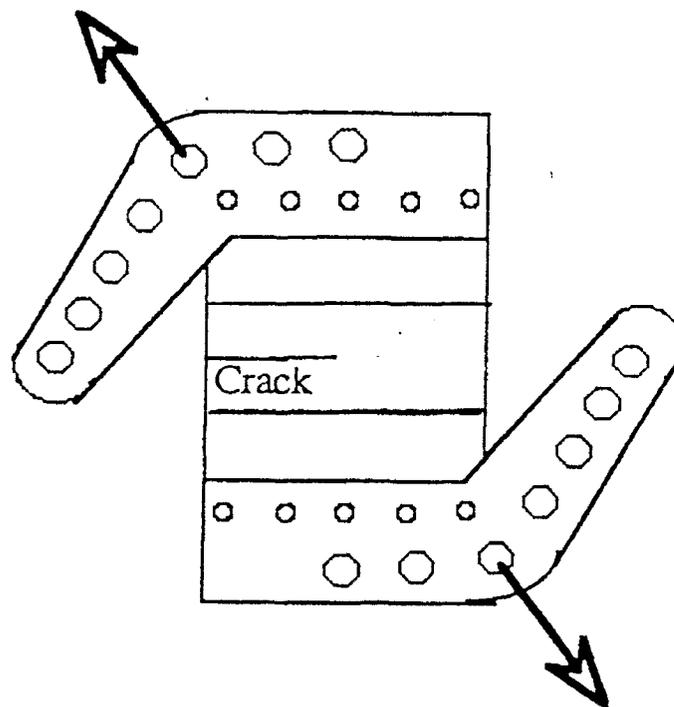


Figure 2. Compact Tension-Shear (CTS) specimen for mixed-mode experiments.

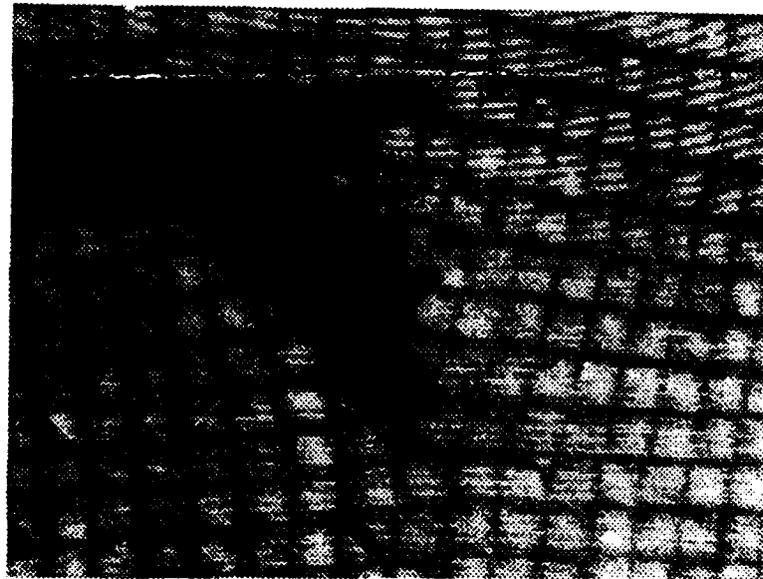
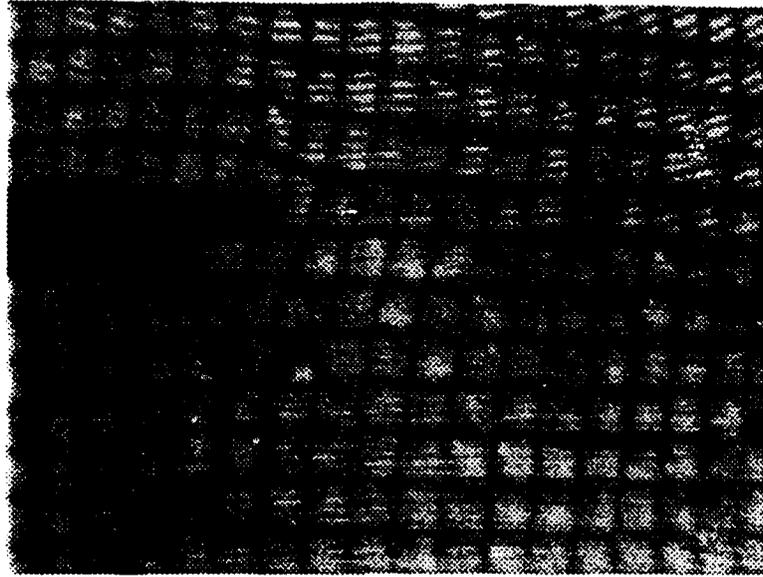


Figure 3. Deformed grid pictures (a) prior to crack initiation and (b) after crack kinking.

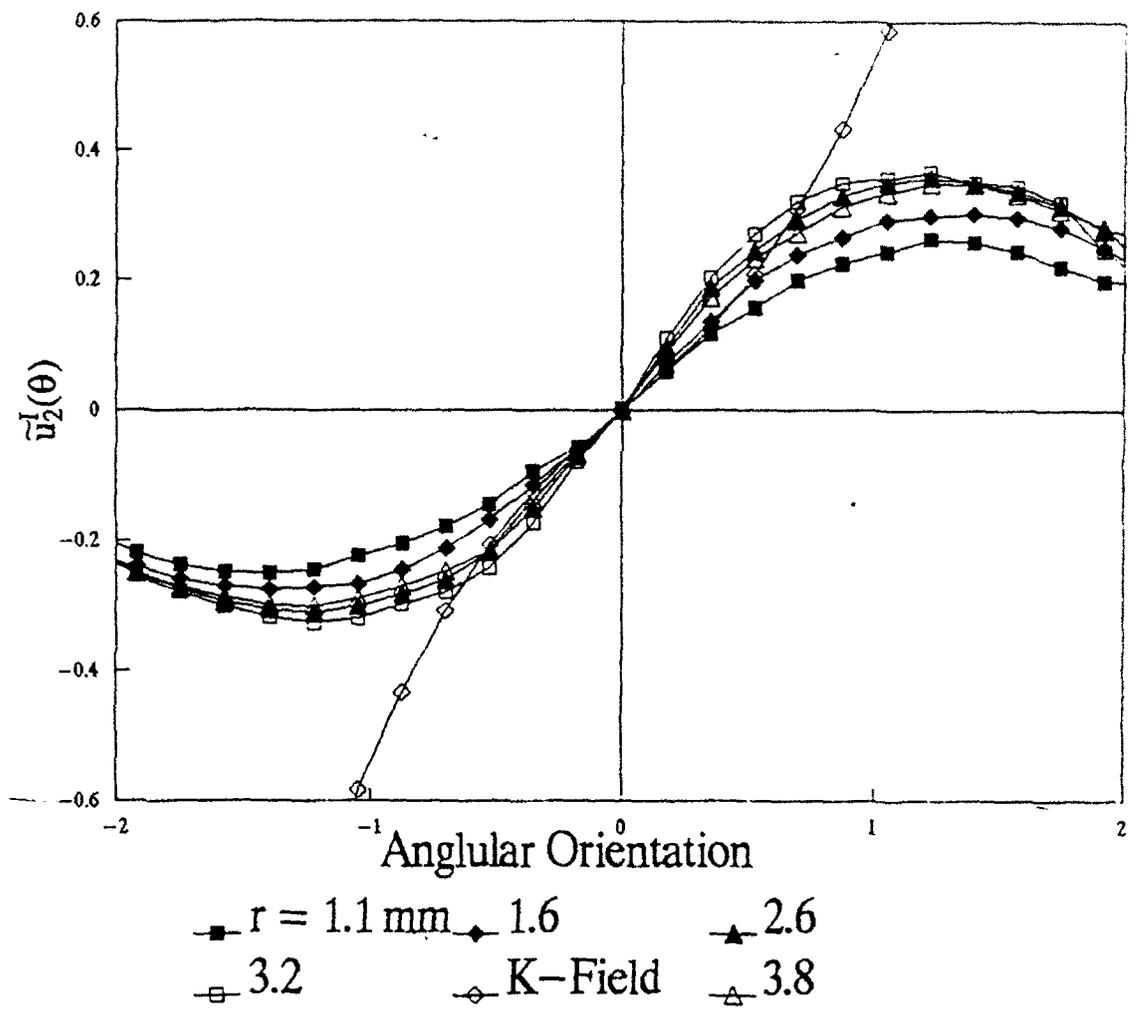


Figure 4. Variation of the crack tip displacement $\bar{u}_2^I(\theta)$ vs θ .

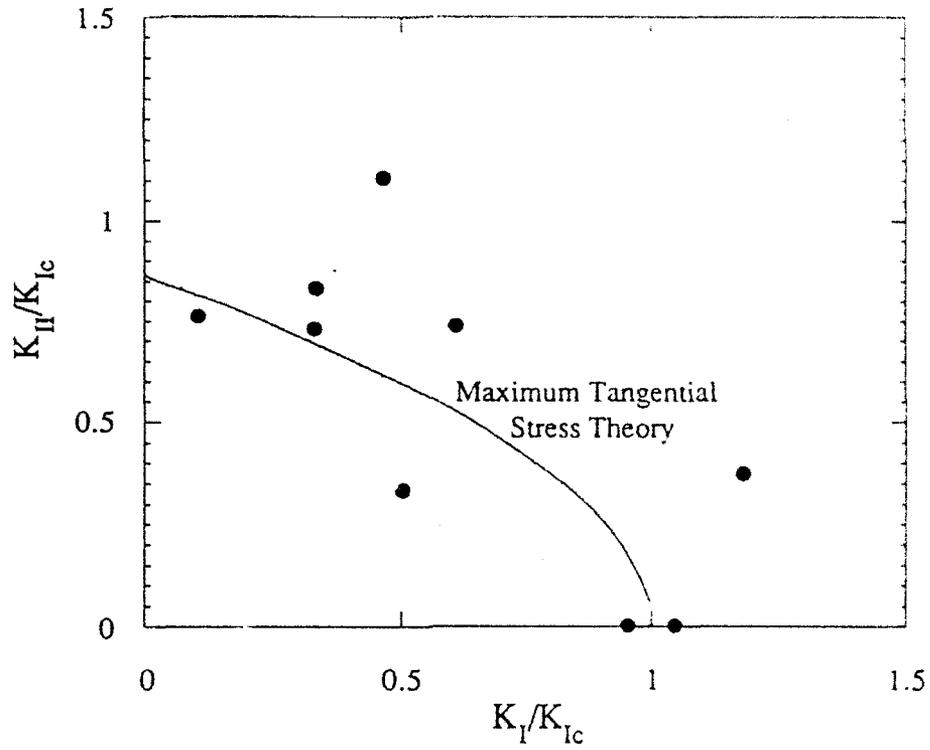


Figure 5. Normalized K_I - K_{II} interaction curve for a solid propellant.

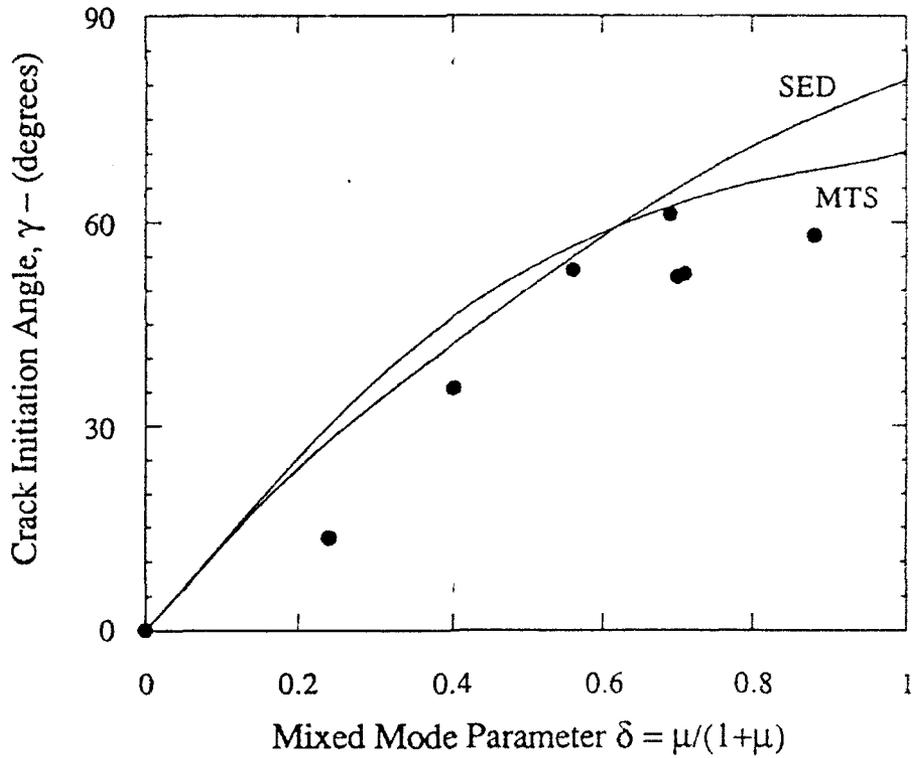


Figure 6. Variation of the crack initiation angle γ with the mixed-mode loading parameter δ for a solid propellant.

VIBRATION AND COMPRESSION TESTING
OF COMPOSITE ISOGRID PANELS

Sean A. Webb
Graduate Student

Dr. Christopher A. Rotz
Associate Professor
Department of Manufacturing Engineering and Engineering Technology

Brigham Young University
435 CTB
Provo, Utah 84602

Final Report for:
AFOSR Summer Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Edwards Air Force Base, California

September 1992

VIBRATION AND COMPRESSION TESTING
OF COMPOSITE ISOGRID PANELS

Sean A. Webb
Graduate Student

Dr. Christopher A. Rotz
Associate Professor
Department of Manufacturing Engineering and Engineering Technology
Brigham Young University

Abstract

Experimental vibration and compression tests of a composite isogrid panel have been successfully carried out. The first four natural frequencies and mode shapes were successfully identified. A simple analytical model developed in the study provided good estimates of the natural frequencies of the bending modes. Bending and torsional modes were analyzed with a simple finite element model. The predicted mode shapes agreed very well with those observed experimentally. Failure loads in the compression test were lower than predicted. This may have been caused by problems with the fixtures used to hold the specimen. Predicted strains were in good agreement with the average values measured in the test. The strains varied more with location on the panel than anticipated in the model, resulting in some strains being higher and others lower than predicted.

VIBRATION AND COMPRESSION TESTING OF COMPOSITE ISOGRID PANELS

Sean A. Webb
Dr. Christopher A. Rotz

1. INTRODUCTION

The Applied Composites Branch at Phillips Laboratory, Space and Missile Technology Directorate, Edwards Air Force Base, California has developed a technique for filament winding fiber reinforced composite isogrid structures. Isogrids would have a shorter lead time for design and manufacture, be faster and less expensive to produce, and be lighter in weight than their aluminum counterparts. Potential areas of application include rocket payload shrouds and fairings, aircraft fuselages, and other aerospace structures.

Detailed computer models of isogrid structures are being developed at Phillips Lab to predict static and dynamic behavior. Presently, the best way to evaluate and compare new isogrid concepts is through mechanical testing. The purposes of this work were: (1) determine the first few natural frequencies and modes of vibration of a flat isogrid panel, (2) test the panel to failure under a static in-plane compressive load, (3) compare the experimental results with those derived from simple analytical and computer models, and (4) make the resulting data available for future comparison with the detailed models under development.

2. ISOGRID PANEL CONSTRUCTION AND DIMENSIONS

The isogrid structures studied in this work were made from T300/976 carbon fiber/epoxy prepreg tow and unidirectional tape. Flat panels were built by hand winding 50 layers of tow into a silicon rubber mold to form the ribs. Since they have all the fibers aligned longitudinally, the ribs would have a modulus along their length $E_r = 21.5$ Mpsi, which is the principal unidirectional modulus of T300/976. The skin was formed by laying the tape over the uncured ribs while they were still in the mold. The layup sequence was $[+60, -60, 0_2, -60, +60]_2$, giving the skin quasi-isotropic in-plane properties. Thus the skin modulus, E_s , equals 8.2 Mpsi, which is the quasi-isotropic modulus for T300/976. The panels were then vacuum bagged and cured for 6 hours at 350°F and 80 psi in an autoclave. After curing, the panels were trimmed to their final size with a diamond saw. A sketch of a finished panel, with all important dimensions, is shown in Fig. 1.

3. VIBRATION TESTING AND ANALYSIS

A series of tests were carried out to determine the first few modes of vibration of an isogrid panel with free (unsupported) edges.

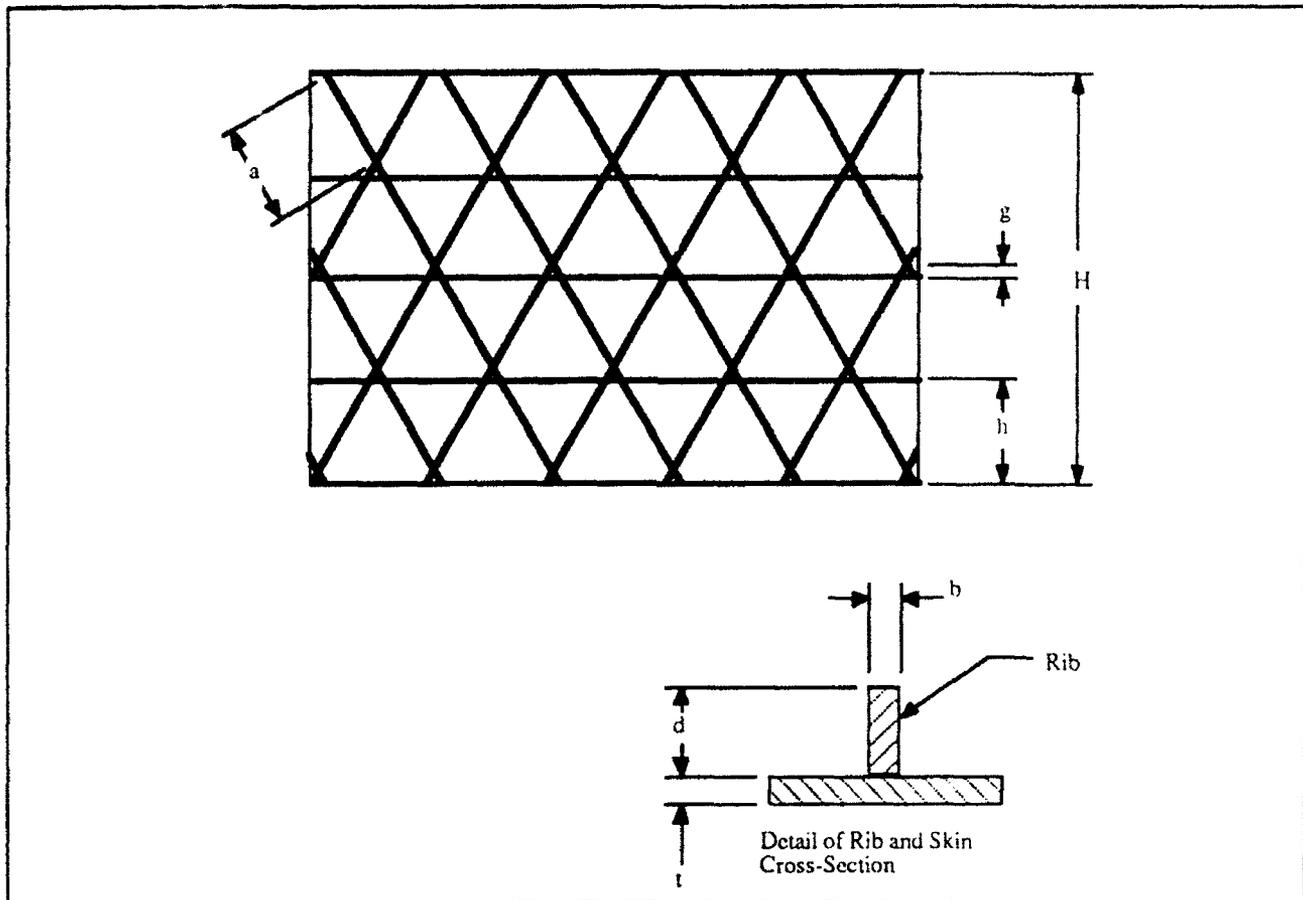


Fig. 1. Sketch of the flat panel isogrid structures studied in this work. The wide black lines are the ribs. The dimensions were: $L = 21.50$ in, $H = 14.13$ in, $d = 0.62$ in, $b = .056$ in, $h = 3.5$ in, $g = 0.4$ in, $t = .053$ in, $a = 3.64$ in. The weight W of the panel was 1.53 lb. The letter P indicates the location of one of the piezoceramic drivers used in vibration testing.

3.1 Apparatus and Procedure

A sketch of the experimental apparatus used in the vibration tests is shown in Fig. 2. Since an instrumented impact hammer was not available, conventional modal testing could not be carried out. Instead the panel was excited by a rectangular shaped piezoceramic element (Type PZT-4, Part #61526, Vernitron Corporation, Bedford, Ohio) bonded to the isogrid structure. In the first series of tests, a 1 in x 2 in x .010 thick piezoceramic driver was mounted on the outer surface of the skin side of the panel ("skin-mounted driver"). The driver was located at the center of the panel with its long sides parallel to those of the panel. In a second series of tests, a 0.5 in x 2 in x .010 in driver was mounted to one of the off-axis ribs ("rib-mounted driver"), as indicated in Fig. 1.

The piezoceramic was driven by swept-sine voltage generated by an HP 35660A Dynamic Signal Analyzer. The voltage was also monitored by one of the analyzer's input channels. The response of the panel was measured with an accelerometer (PCB SN 6471 336B) whose output

was amplified (PCB 480 D06 Charge Amplifier) before being fed into the second channel of the signal analyzer.

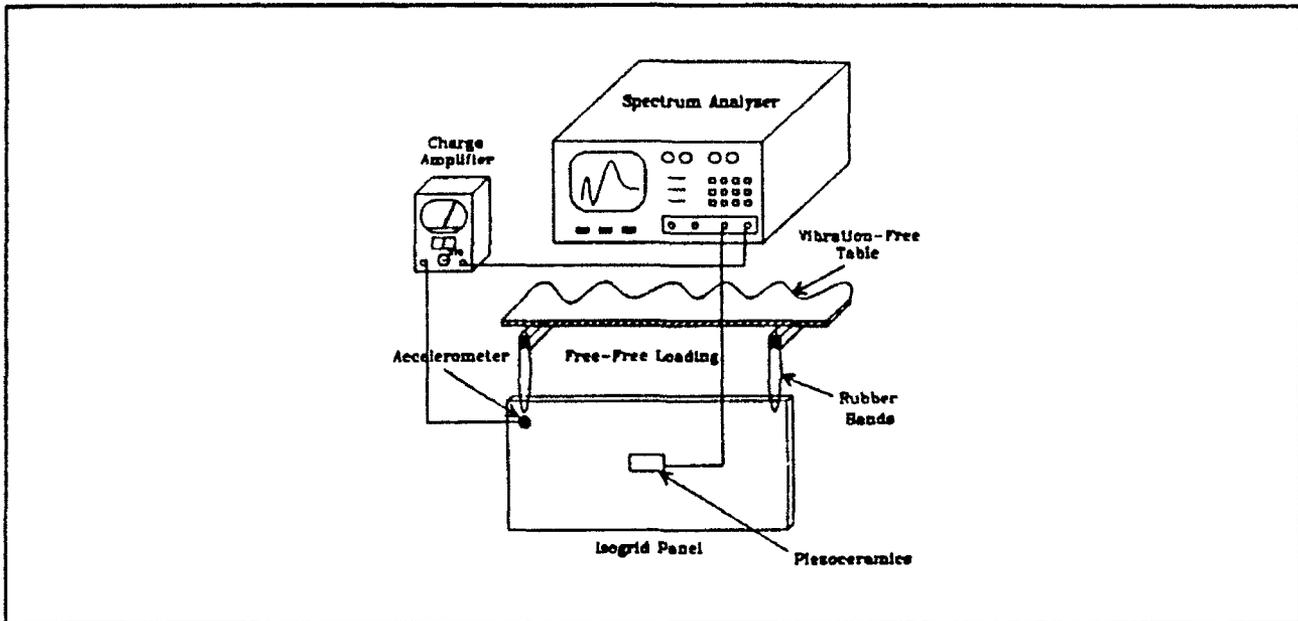


Fig. 2. Schematic of the experimental apparatus used to measure the natural frequencies and mode shapes of the isogrid panel. The sketch shows the configuration for the skin-mounted piezoceramic driver. The rib-mounted driver is not visible here.

A grid was laid out on the skin surface of the panel, and the accelerometer was systematically placed at various locations to map out the response. The analyzer calculated the transfer function between the voltage input and the acceleration response (10 averages, uniform window) at each location on the panel. All accelerations were measured normal to the skin surface.

The magnitude and phase of the transfer functions at various resonance frequencies were used to plot approximate mode shapes. The phase data were only used to determine whether the response at each node was in phase or out of phase with the input. Such an approximation is valid for lightly damped structures like the ones involved here.

3.2 Experimental Results

A typical transfer function measurement for the skin-mounted driver is shown in Fig. 3. Resonance peaks can be observed at 296, 536, and 844 Hz. An additional resonance at 232 Hz was observed when the rib-mounted driver was used. There is one other peak visible at about 560 Hz. Since it is so close to the one at 536 Hz, it would be difficult if not impossible to extract the associated mode without more sophisticated equipment than was available to the authors. No attempts were made to differentiate between the two modes. The 560 Hz peak was only observed in a few of the transfer function measurements.

The measured mode shapes for the four frequencies (232, 296, 536, and 844 Hz) are shown in Figs. 4-11. The mode at 232 Hz, shown in Figs. 4 and 5, appears to be a torsional

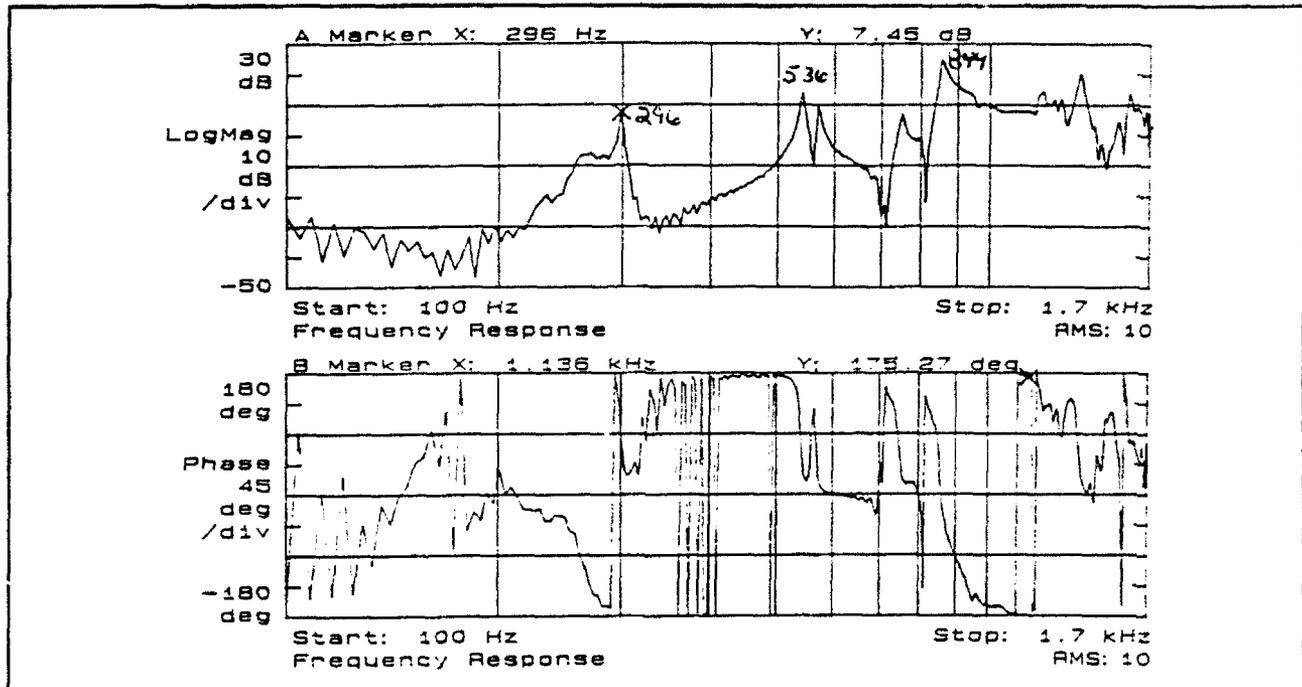


Fig. 3. Typical results for the transfer function between the input voltage and the acceleration response at one point on the skin of the panel. The above results were obtained with a skin-mounted piezoceramic driver. Only the rib-mounted driver excited a resonance at 232 Hz.

mode. The mode at 296 Hz, shown in Figs. 6 and 7, involves bending about an axis parallel to the short sides of the panel, and will be called the lateral bending mode. The mode at 536 Hz, shown in Figs. 8 and 9, involves bending about an axis parallel to the long sides of the panel, and will be called the longitudinal bending mode. The results for the mode at 844 Hz, shown in Figs. 10 and 11, are too irregular to interpret. It may be that the acceleration signals are so weak at this high frequency that noise has obscured the measurements.

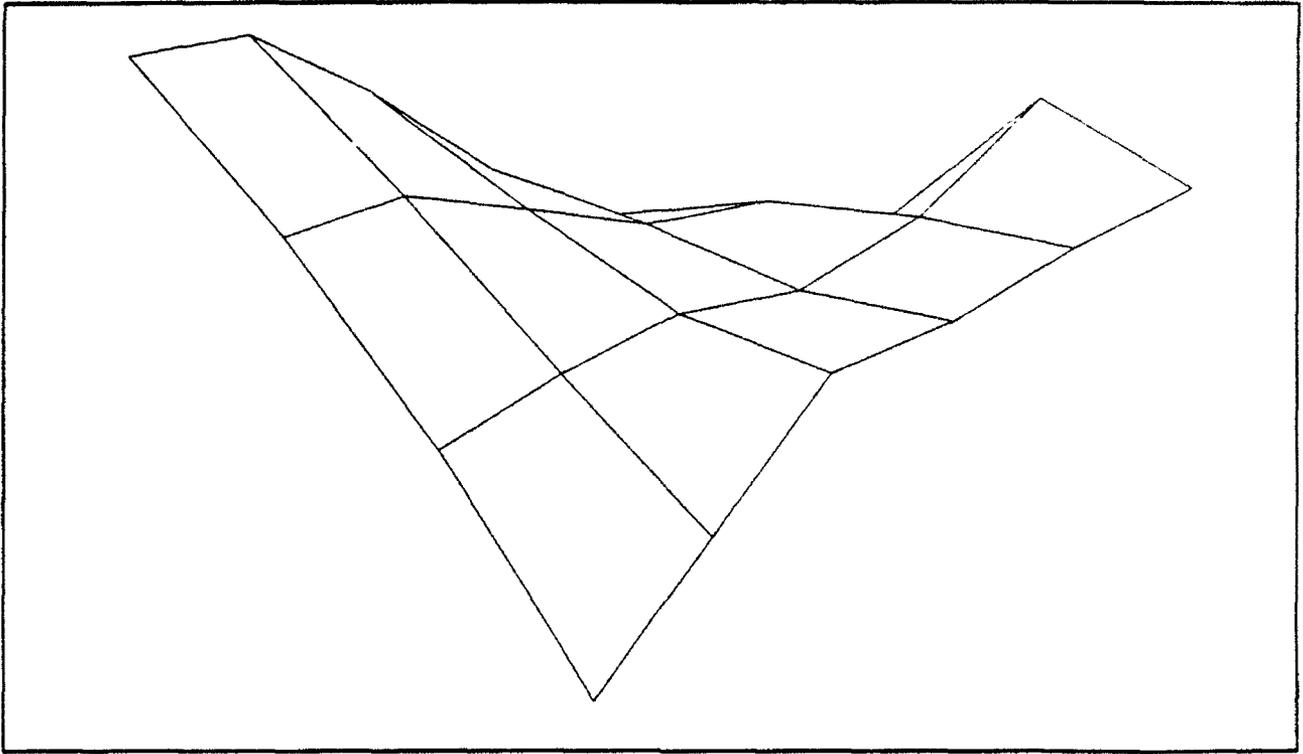


Fig. 4 Experimental results for the first torsional mode (232 Hz): deformed shape plot.

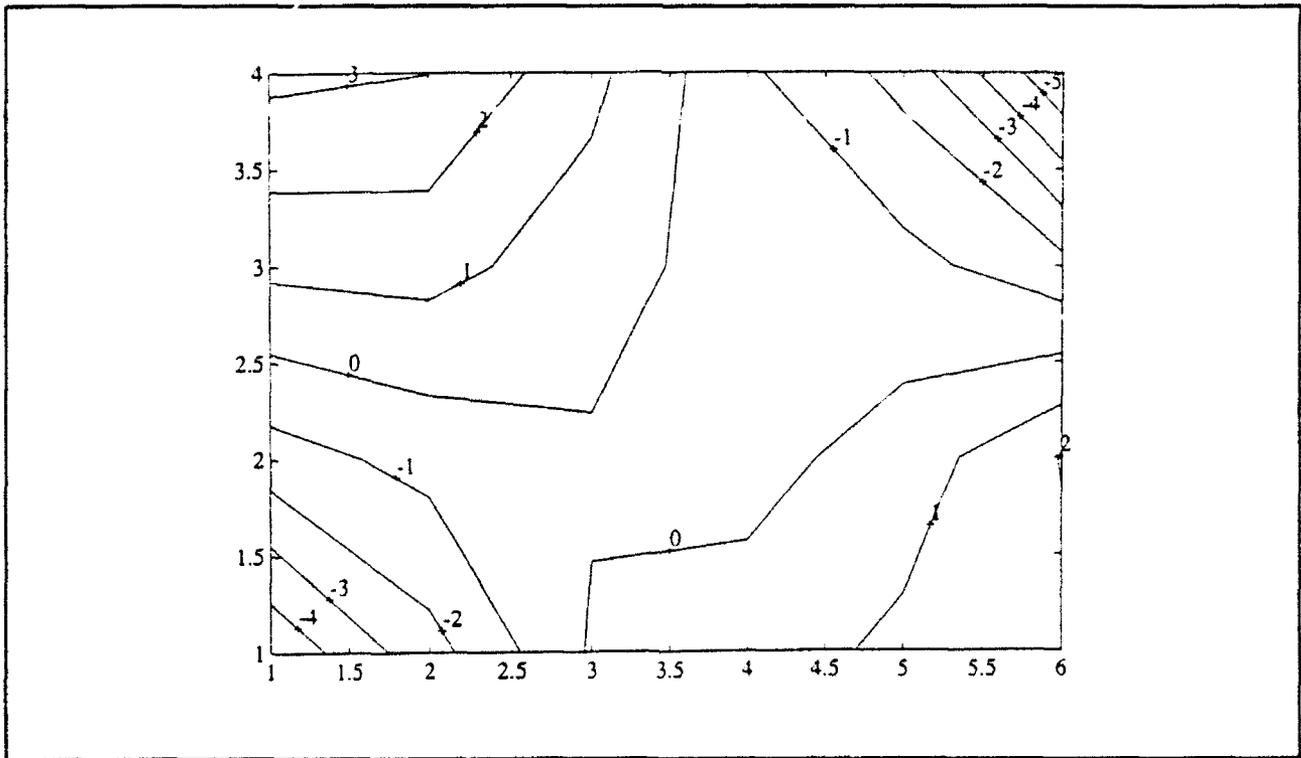


Fig. 5. Experimental results for the first torsional mode (232 Hz): contour plot.

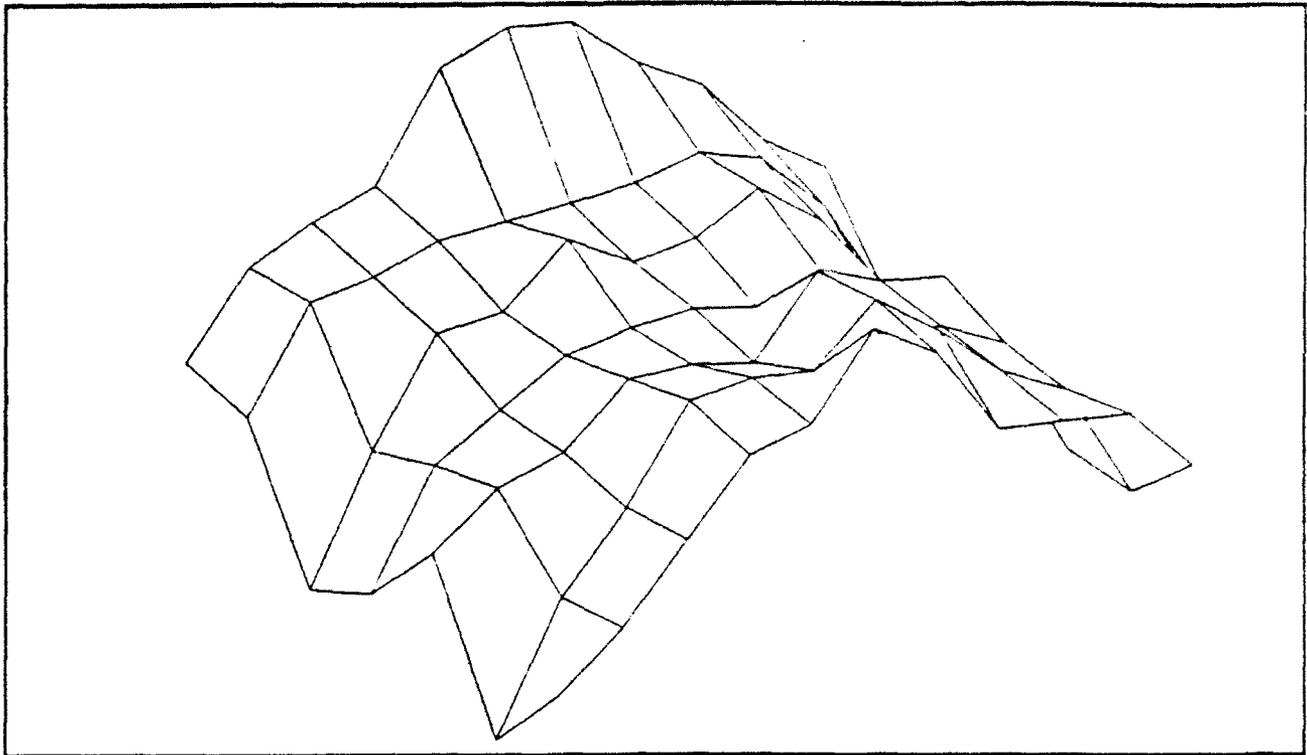


Fig. 6. Experimental results for the first lateral bending mode (296 Hz): deformed shape plot.

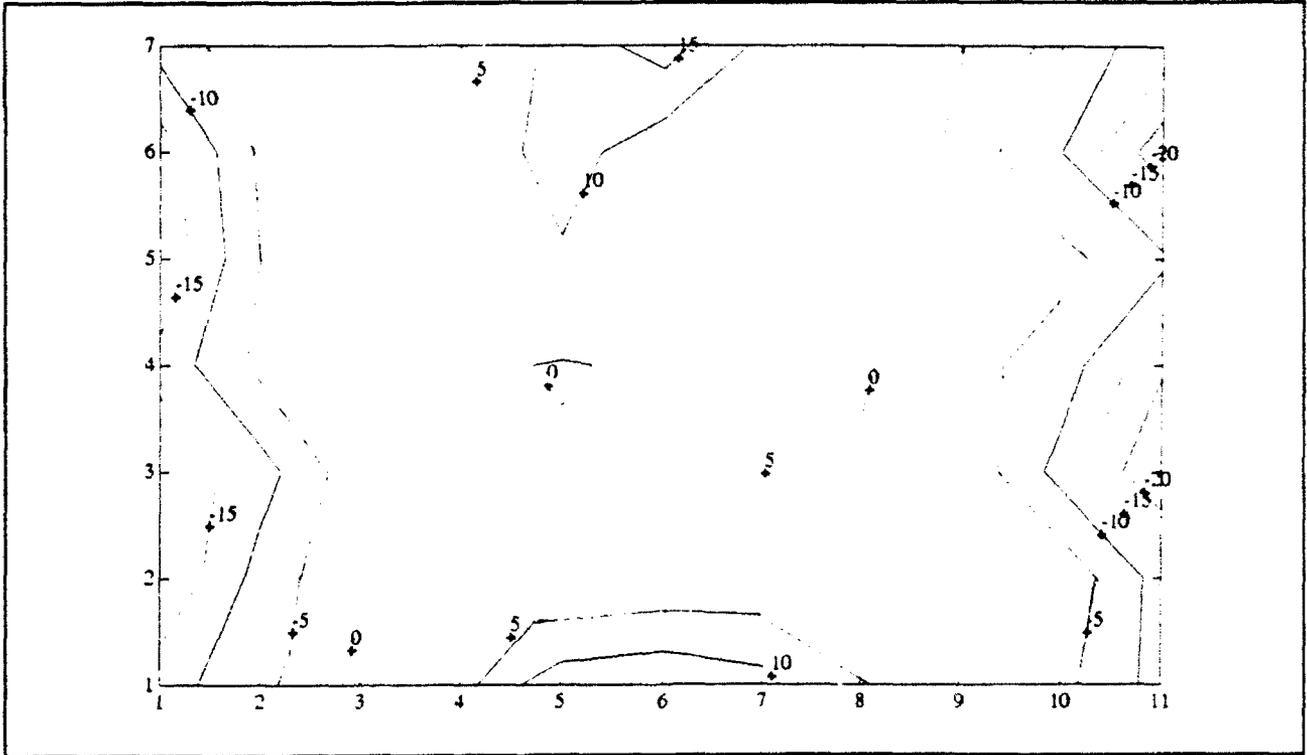


Fig. 7. Experimental results for the first lateral bending mode (296 Hz): contour plot.

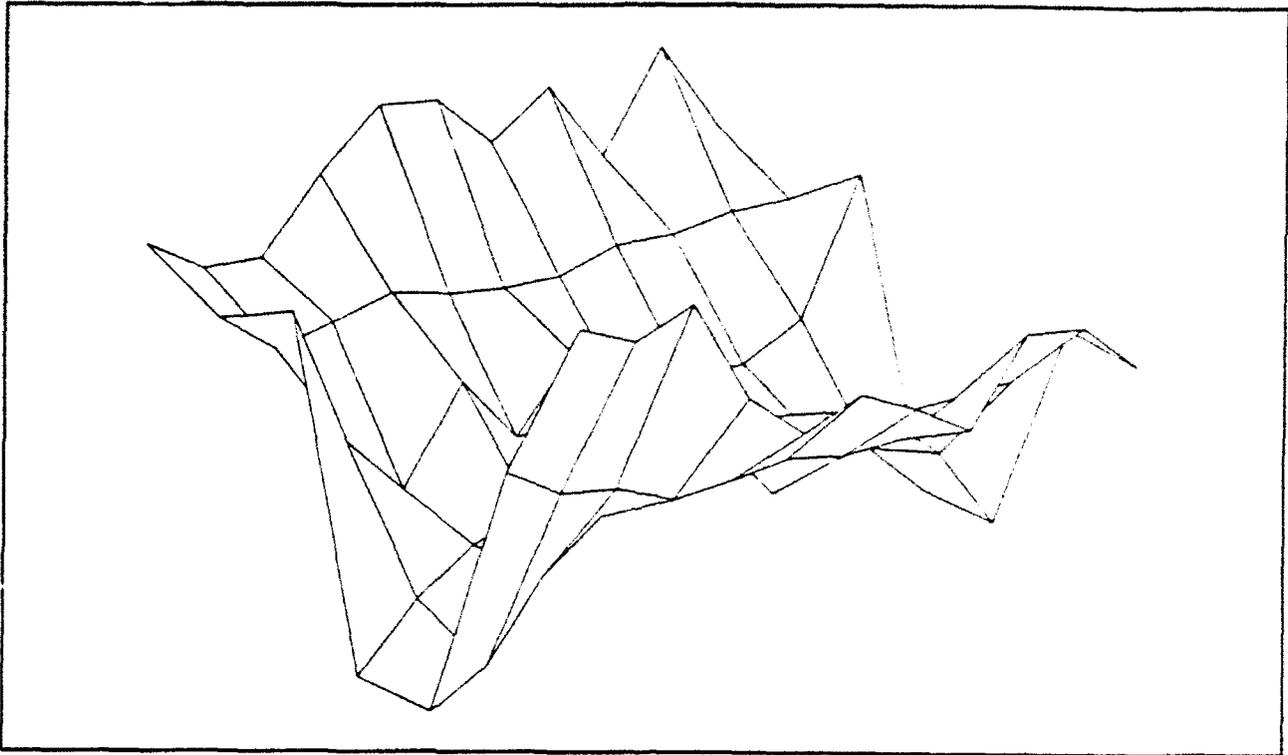


Fig. 8. Experimental results for the first longitudinal bending mode (536 Hz): deformed shape.

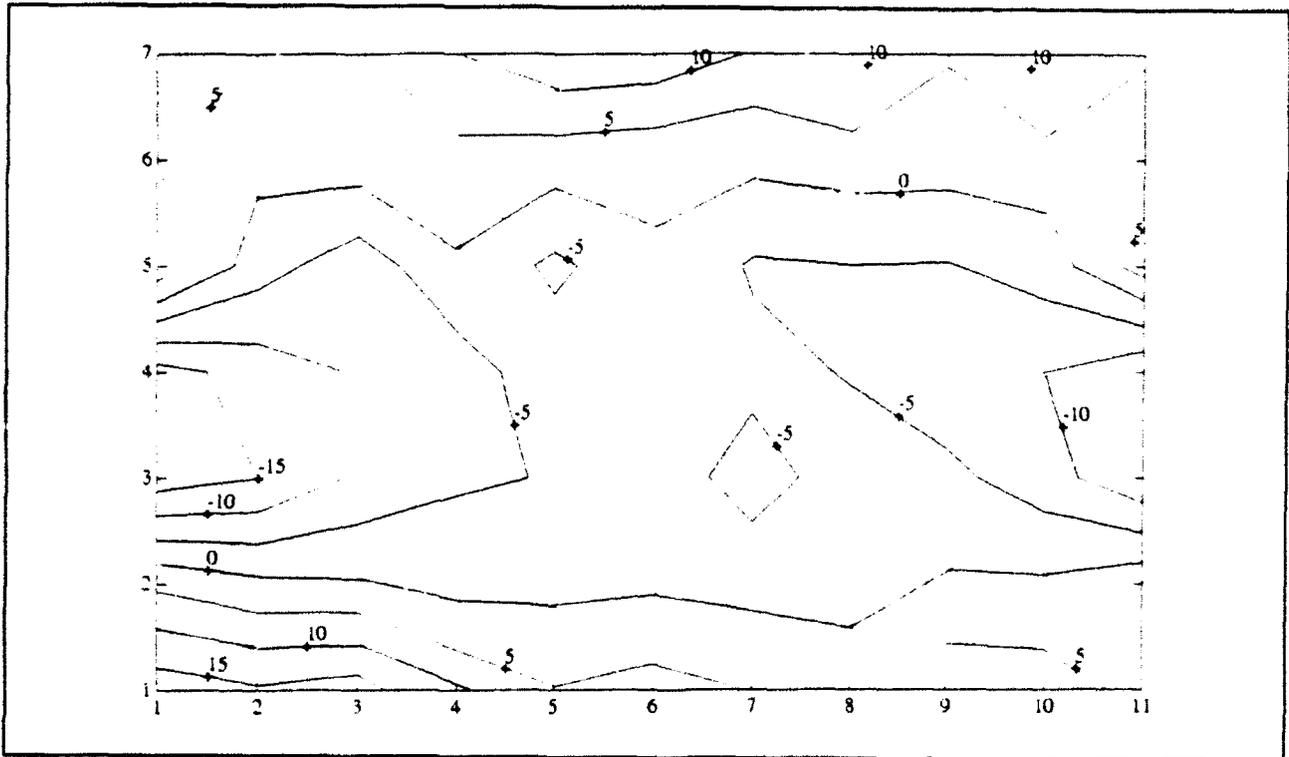


Fig. 9. Experimental results for the first longitudinal bending mode (536 Hz): contour plot.

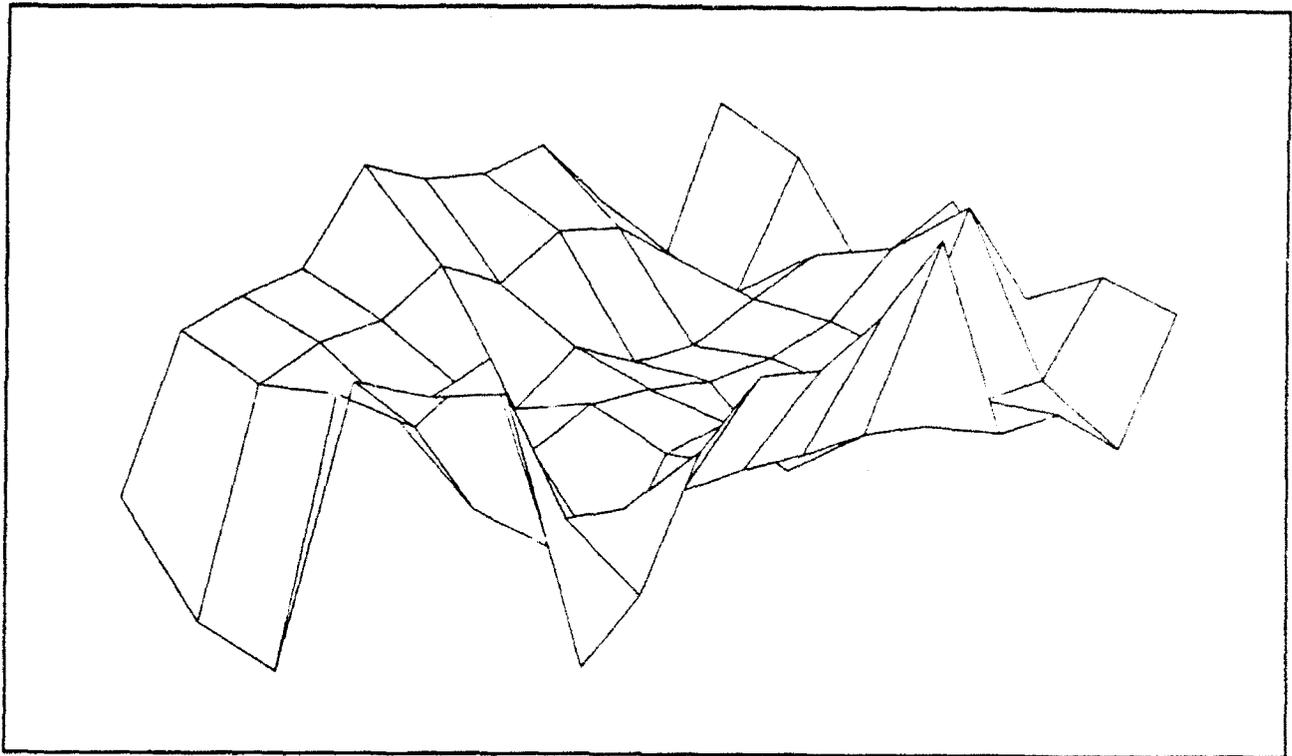


Fig. 10. Experimental results for the mode at 844 Hz: deformed shape plot.

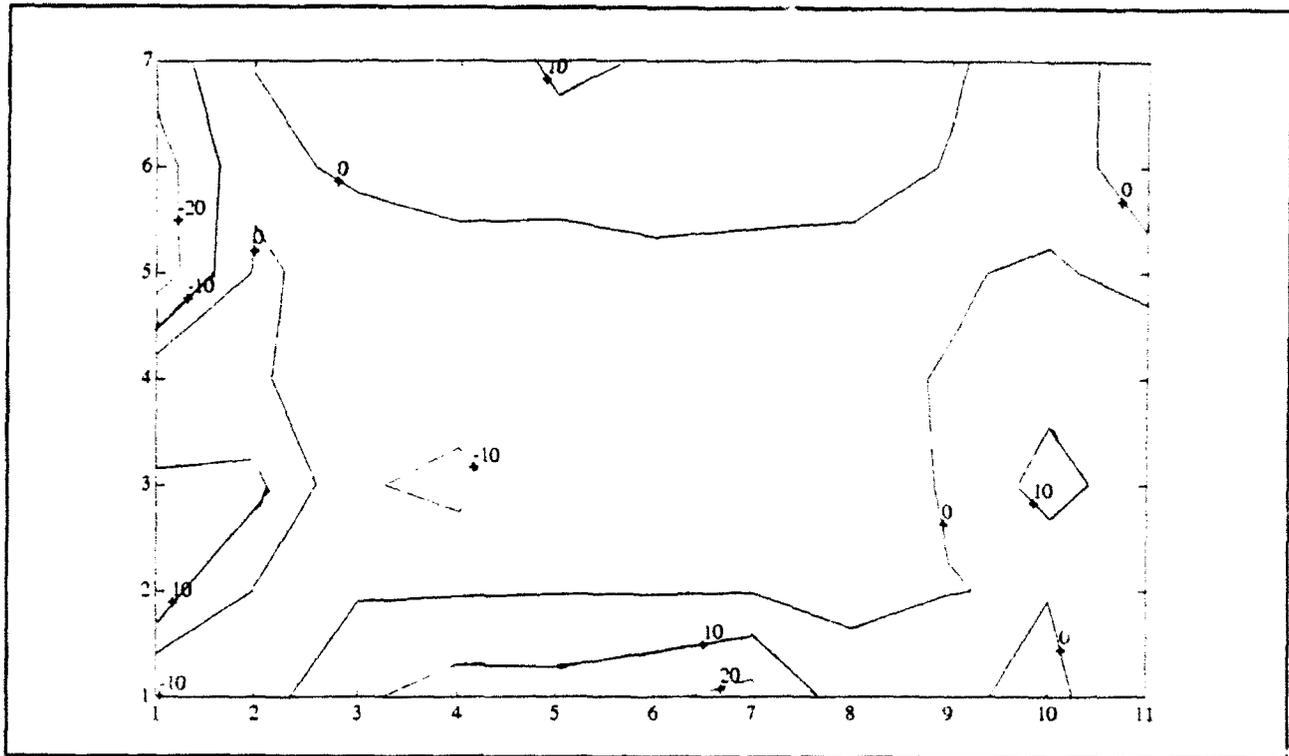


Fig. 11. Experimental results for the mode at 844 Hz: contour plot.

3.3 Analytical Predictions

The isogrid panel can be approximately modeled as a rectangular plate of length L , width H , equivalent thickness t^* , equivalent modulus E^* , and weight W . (Such an approximation is sometimes referred to as a "smeared" model.) It can be shown that

$$t^* = \sqrt{\frac{12 I_{\text{eff}}}{t_a}} \quad \text{and} \quad E^* = \frac{t_a}{t^*} E_s \quad (1), (2)$$

where

$$t_a = t(1 + \alpha) \quad I_{\text{eff}} = \frac{t^3 \beta^3}{12(1 + \alpha)} \quad \alpha = \frac{E_r}{E_s} \frac{bd}{ht} \quad (3a,b,c)$$

$$\beta^2 = (1 + \alpha\delta^2)(1 + \alpha) + 3\alpha(1 + \delta)^2 \quad \delta = \frac{d}{t} \quad (3d,e)$$

In the above equations, E_s is the modulus of the isogrid skin (8.2 Mpsi) and E_r is the modulus of a rib along its length (21.5 Mpsi). Note that b , d , t , h , L and H are defined in Fig. 1.

3.3.1 Analytical Predictions of the Natural Frequencies for the Panel Bending Modes

No exact analytical solutions exist for the vibration of a freely suspended plate. However, the bending modes can be approximated by the modes of a free-free beam. For a beam of length L_b , area moment of inertia I_b , weight W_b , and modulus E_b , the first two natural frequencies, ω_1 and ω_2 , are:

$$\omega_1 = 22.4 \sqrt{\frac{E_b I_b}{W_b L_b^3}} \quad \omega_2 = 61.7 \sqrt{\frac{E_b I_b}{W_b L_b^3}} \quad (4a,b)$$

The natural frequencies of the first two lateral bending modes of the isogrid panel can be estimated from Eqs. 4a and 4b by substituting

$$L_b = L \quad I_b = \frac{H t^{*3}}{12} \quad E_b = E^* \quad W_b = W \quad (5a,b,c,d)$$

Similarly, the natural frequency of the first longitudinal bending mode can be estimated from Eq. 4a by substituting

$$L_b = H \quad I_b = \frac{L t^{*3}}{12} \quad E_b = E^* \quad W_b = W \quad (6a,b,c,d)$$

No torsional modes can be estimated with this approach. The results for the isogrid panel studied in this work are summarized in Table 1.

3.3.2 Approximate Finite Element Analysis for the Natural Frequencies and Modes

The natural frequencies and mode shapes of all the modes of vibration were estimated using the IDEAS finite element analysis program. The isogrid was modeled as a uniform flat plate of length L, width W, thickness t*, weight W, and modulus E*. The free edge conditions were approximated by supporting corners of the plate with spring elements having very low stiffness. Such an analysis was carried out using the IDEAS finite element analysis package. Both mode shapes and natural frequencies were determined. The results are summarized in Table 1.

Table 1. Predicted and observed natural frequencies for the isogrid panel.

Mode Description	Free-Free Beam Approximation	Approximate Finite Element Model	Experimental Results
1st Torsion	-----	305 Hz	232 Hz
1st Lateral Bending	326 Hz	323 Hz	296 Hz
2nd Torsion	-----	694 Hz	-----
1st Longitudinal Bending	755 Hz	756 Hz	536 Hz
2nd Lateral Bending	898 Hz	869 Hz	844 Hz?
?	-----	1006 Hz	-----

3.4 Discussion of Results

Bending Modes: The approximation based on a free-free beam model accurately predicted all the observed natural frequencies for the bending modes. However it must be noted that observed mode at 844 Hz may or may not be bending; the tests were inconclusive for the reasons discussed earlier. The computer model predicted shapes that agreed well with the test results for the first and second lateral bending, and the first longitudinal bending modes. The computer predicted a second lateral bending mode at 869 Hz which could not be verified. The frequencies predicted for the lateral bending modes agree well with experiment, but the one for the longitudinal mode differs by 25-30%.

Torsional Modes: The computer model accurately predicted the mode shape for the first torsional mode. The predicted frequency was about 30% too high. What appeared to be a second torsional mode at 694 Hz was also predicted, but not found experimentally.

4. COMPRESSION TESTING AND ANALYSIS

The following section will describe the in-plane compression testing of the isogrid panel. The tests were run to accumulate data for future comparison with detailed finite element models of the isogrid.

4.1 Apparatus

Strain gages (type CEA-06-250UN-350, Micromeritics Group, Inc., Raleigh, NC) were attached to the panel at the locations shown in Fig. 12. The gages were connected in a three-wire configuration to a strain gage energizing and signal conditioning system. The excitation voltage for each bridge circuit was set at $2.000 \pm .001$ v. The conditioned signals were fed into a Macintosh IIx based data acquisition system capable of monitoring each channel with 16 bit precision.

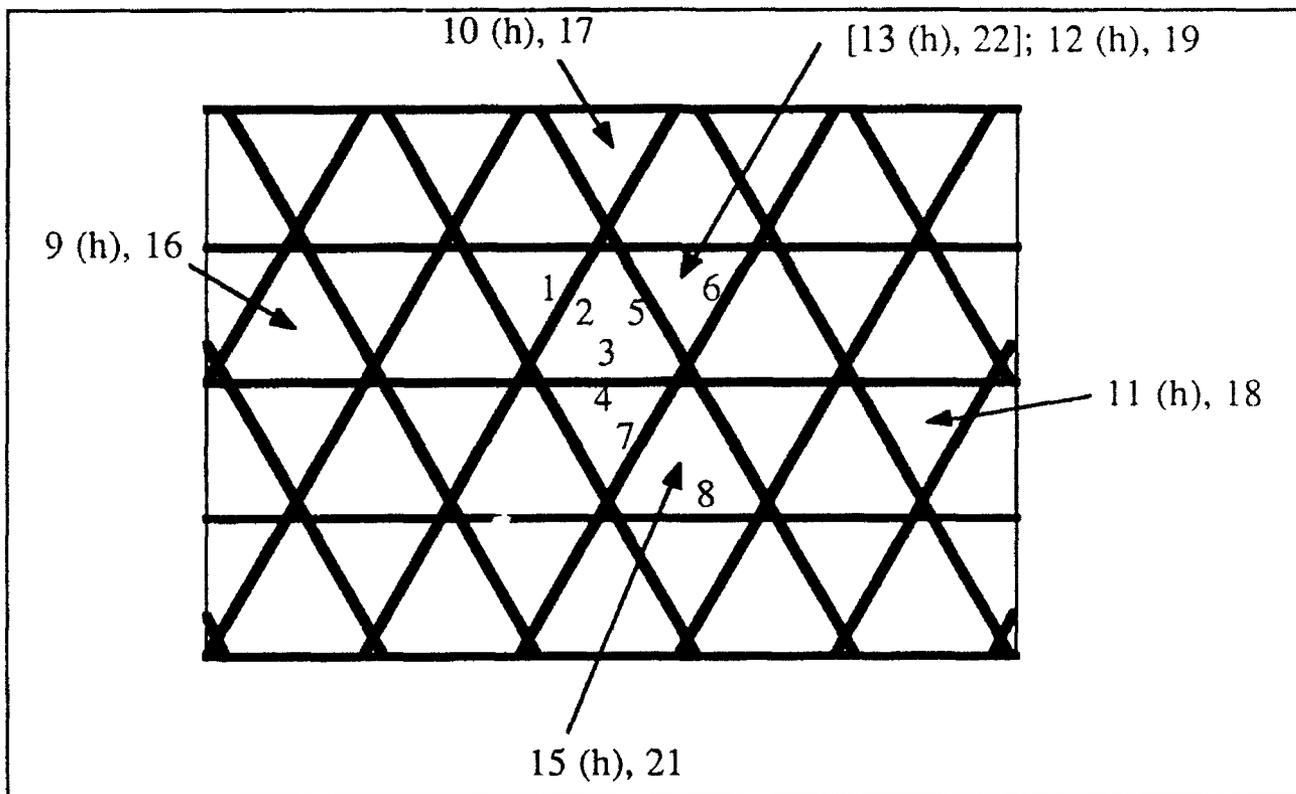


Fig. 12. Schematic diagram of strain gage placement for panel compression test. Gages attached to the ribs are indicated by the numbers next to the ribs in the figure. All such gages were aligned with the ribs. Gages attached to the skin are indicated by the arrows. These gages were mounted in the center of each triangular pocket. Horizontally oriented gages are indicated by an "(h)" following the number. All other gages were mounted vertically. All gages were mounted on the rib side of the panel except the ones whose numbers are enclosed in square brackets.

The panel was clamped in a special aluminum fixture sketched in Fig. 13. The fixtures enabled the panel to be loaded in compression on an MTS servo-hydraulic tensile test machine. Voltage signals from the load cell and crosshead displacement sensor were fed into the computer through two additional channels of the previously described data acquisition system.

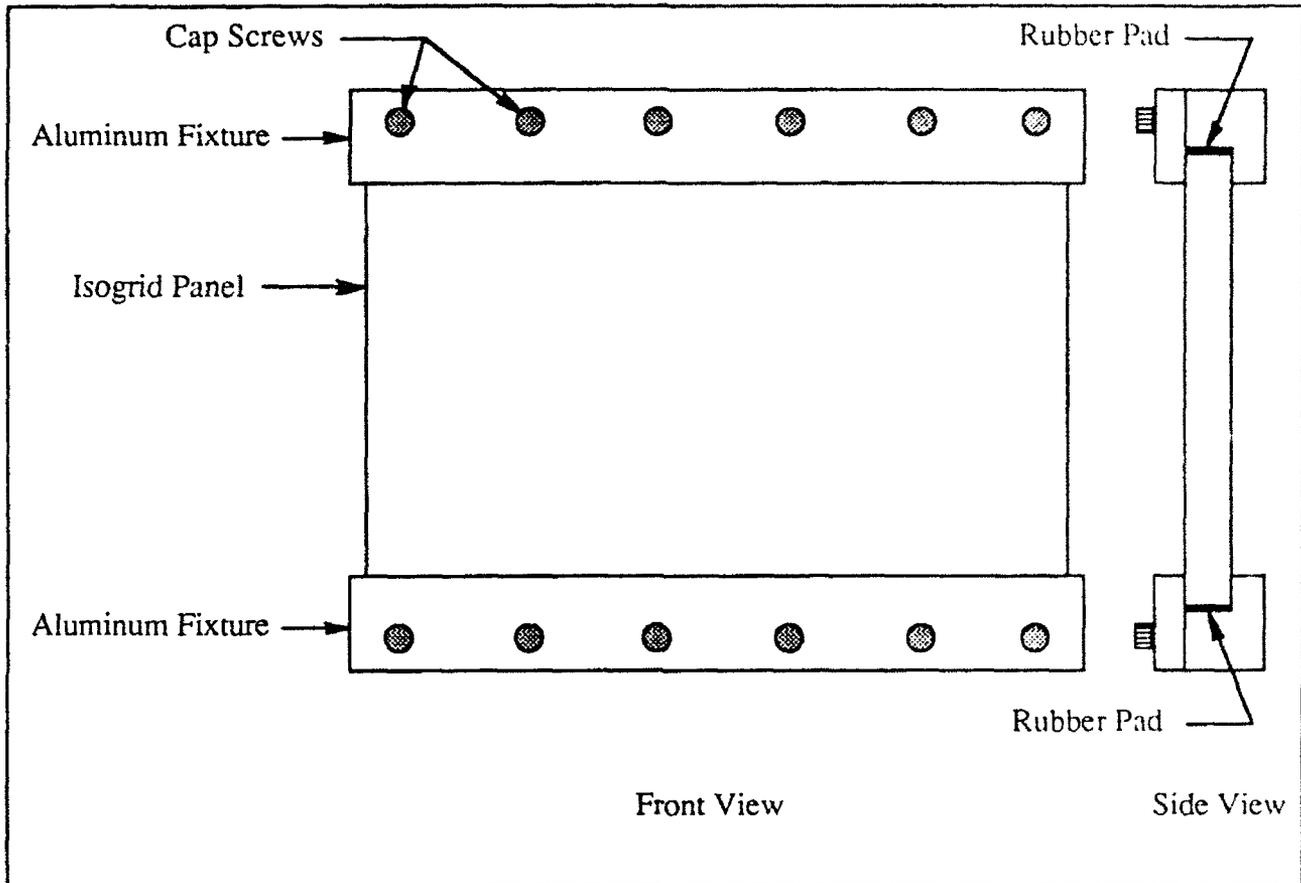


Fig. 13. Schematic diagram of panel mounting scheme for the compression test. (The isogrid ribs have been left out of the sketch for simplicity.) The aluminum fixtures were machined so that the panel was firmly clamped in place when the cap screws were tightened. The load acted in a vertical direction. The 0.1 in thick rubber pads were used to distribute the load as uniformly as possible along the length.

4.2 Procedure

After the gages were attached to the panel, they were checked with a gage installation tester (Micromeritics Model 1300). The panel was then mounted in the MTS test machine, and the crosshead was carefully lowered to bring it into contact with the upper fixture.

The wires from the gages, load cell, and crosshead displacement sensor were connected to the data acquisition system. The input voltages to each bridge circuit were set, and the circuits individually balanced. Each circuit was then shunt calibrated with a known resistance. The calibration factor was then entered into the computer for each channel. Similarly, the calibration

factors for the load cell and displacement sensor were entered. The values were obtained from the operator's manual for the MTS machine. The values were checked using a calibration procedure outlined in the manual. The sampling rate was entered into the computer (a value of 1 Hz was used in all cases), and the test begun.

4.3 Experimental Results

The compressive load is plotted as a function of vertical displacement in Fig. 14. The measured strains are plotted as functions of the applied compressive load in Figs. 15-17. The panel failed at a maximum compressive load of 6783 lb. Failure appeared to be initiated by buckling of the ribs at a load near 6000 lb. Measured strain values became very erratic for loads above 6000 lb.

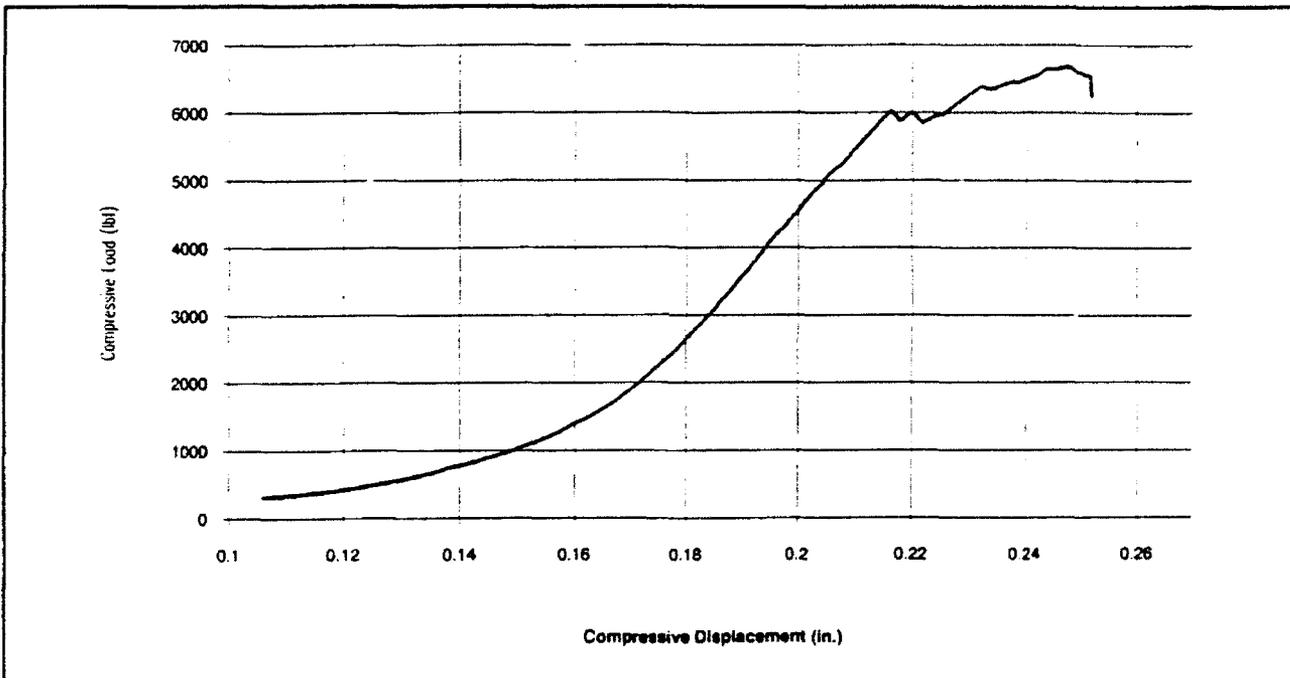


Fig. 14. Compressive load on the panel as a function of crosshead displacement.

T300/976 Isogrid Compression Test (Rib Gages)

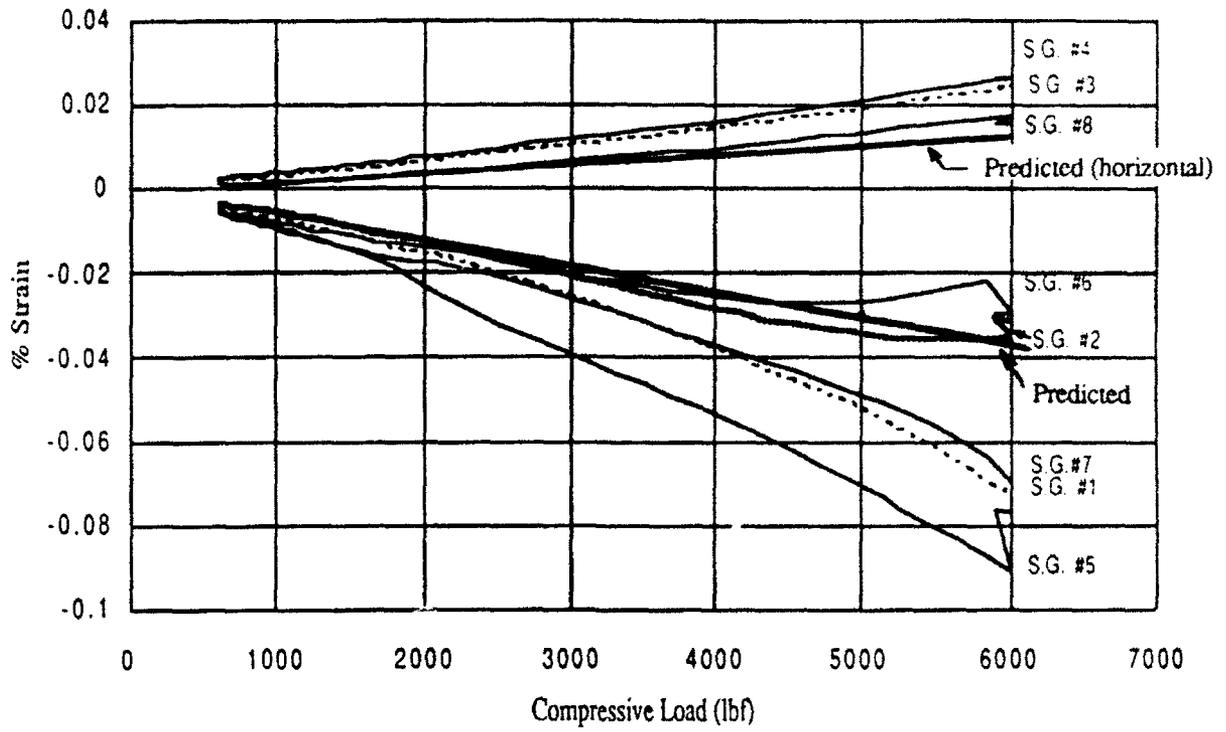


Fig. 15. Measured strains in the ribs of the panel.

T300/976 Isogrid Compression Test (Horizontal Skin Gages)

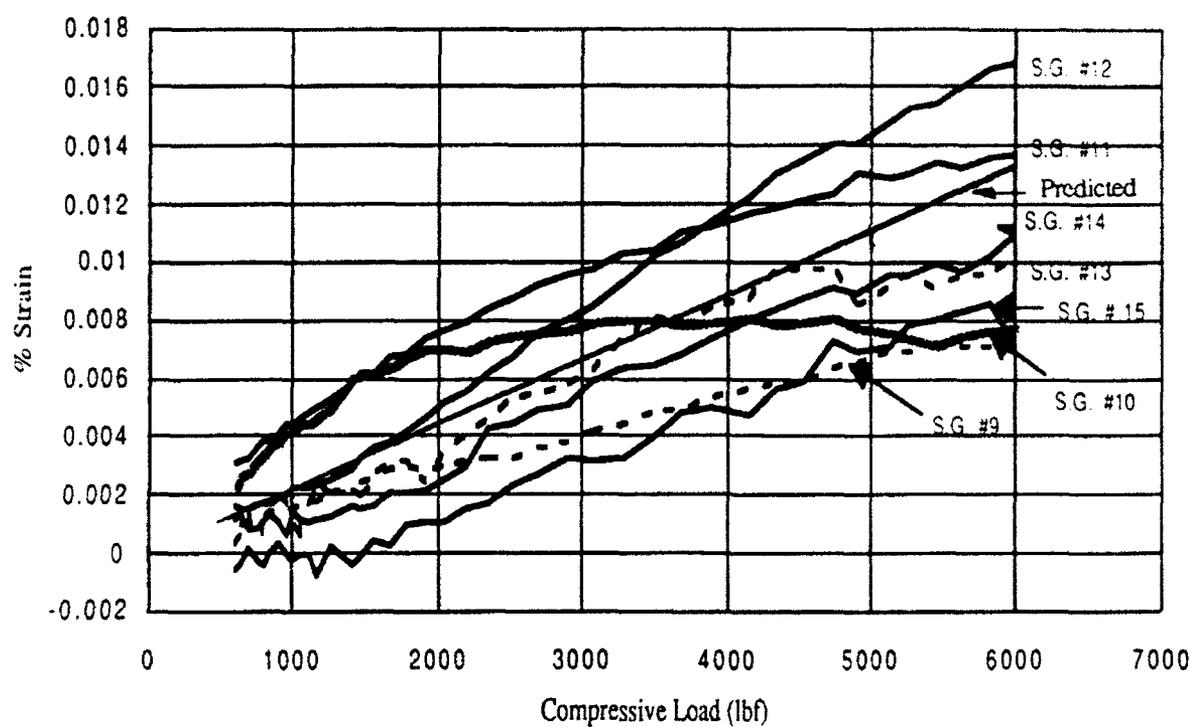


Fig. 16. Measured horizontal strains in the skin of the panel.

T300/976 Isogrid Compression Test (Vertical Skin Gages)

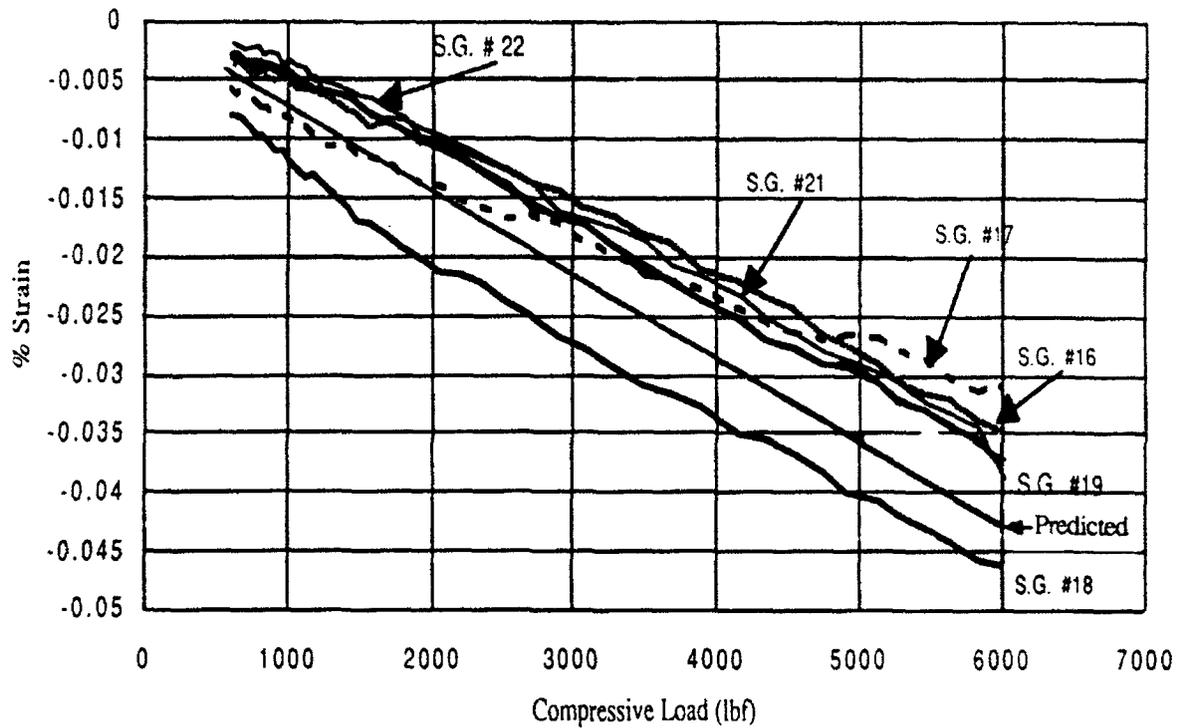


Fig. 17. Measured vertical strains in the skin of the panel.

4.4 Theoretical Predictions

Predictions of the effects of an in-plane compressive load F can be made by first modeling the isogrid as a uniform plate of length L , width H , thickness t_{eff} , and modulus E_r (the longitudinal modulus of the ribs), where

$$t_{eff} = t \left(\frac{E_s}{E_r} + \frac{b d}{t h} \right) \quad (7)$$

The strains in the vertical (y) and horizontal (x) directions can then be calculated for the plate. Since the skin and ribs must move together, the predicted horizontal strain ϵ_{hs} in the skin would be equal to the strain in the horizontal ribs ϵ_{hr} and both would be equal to ϵ_x :

$$\epsilon_{hs} = \epsilon_{hr} = \epsilon_x = \frac{F \nu}{E_r L t_{eff}} = 22 \frac{\text{microstrain}}{\text{kip}} F \quad (8)$$

where ν is the Poisson's ratio of the composite material (about 0.32).

The predicted vertical strain in the skin ϵ_{vs} would be equal to ϵ_y :

$$\epsilon_{vs} = \epsilon_y = \frac{-F}{E_r L_{teff}} = -72 \frac{\text{microstrain}}{\text{kip}} F \quad (9)$$

The strain in the inclined ribs, ϵ_{ir} , can be found using a Mohr's circle analysis to determine the strain components along the ribs. The results are

$$\epsilon_{ir} = -(0.25 \nu + 0.75) \frac{F}{E_r L_{teff}} = -60 \frac{\text{microstrain}}{\text{kip}} F \quad (10)$$

The predicted strains are plotted in Figs. 15-17.

An analysis to predict the failure load on the panel was carried out. It was assumed that the panel would fail either by buckling of the ribs or by buckling of the triangular regions (pockets) enclosed by the ribs. The failure load for rib buckling F_{rb} was estimated to be:

$$F_{rb} = \left(\frac{C \pi^2}{12 E_r (1 - \nu_x \nu_y)} \left(\frac{b}{a} \right)^2 + G_{xy} \left(\frac{b}{d} \right)^2 \right) L t \left(\frac{E_s}{E_r} + \frac{b d}{h t} \right) = 12000 \text{ lb to } 27000 \text{ lb} \quad (11)$$

where $C = 1$ if the ribs are assumed to be simply-supported at the vertices of the triangles in the isogrid, or $C = 4$ if the ends are assumed to be fixed. In Eq. 11, ν_x and ν_y are the in-plane Poisson's ratios (0.32 and .015) and G_{xy} the shear modulus of the rib material (0.5 Mpsi), and a is the length of ribs between the nodes (see Fig. 1).

The predicted load for pocket buckling, F_{pb} , is given by the equation

$$F_{pb} = \left(k_1 \frac{11.0 \pi^2 E_s}{12 (1 - \nu^2)} \right) \left(\frac{t}{h} \right)^2 L t \left(\frac{E_s}{E_r} + \frac{b d}{h t} \right) = 67000 \text{ lb} \quad (12)$$

where k_1 is a constant which depends on the skin layup (approximately = 1). Thus according to this simple analysis failure by rib buckling would be expected.

4.4 Discussion of Compression Test Results

The lowest predicted failure load (12000 lb) was about twice the actual failure load (6000-6785 lb). Perhaps the aluminum fixtures used to hold the panel do not give adequate support. Any small angular flexibility at the supports will produce a bending moment in the panel, thus causing an apparent premature failure. It is recommended that, in the future, the panels should be "potted" into the fixtures with an epoxy or similar material. Secondly, it is apparent from the strain plots that the load was not uniformly distributed throughout the panel. Part of this is to be expected since the lateral edges are stress free. However, the differences are not symmetrically distributed about the panel center line leading to the conclusion that the load was not evenly transmitted along the length of the panel. Perhaps the rubber pads are insufficient to accommodate

variations in the panel geometry. Again, potting the ends with epoxy could overcome this problem.

The predicted strains in the ribs are close to the actual measured values for some of the gages, and smaller in magnitude for the others. The predicted strains in the vertical direction for the skin are larger in magnitude than the measured values by about 20% for all the gages except one. The predicted strains in the horizontal direction in the skin are in the middle of the measured values. Given the spread in the measured strain values (due to the non-uniform load distribution and/or edge effects), the simple analytical models seem to provide at least a semi-quantitative prediction of the actual behavior.

5. CONCLUSIONS

Experimental vibration and compression tests of a composite isogrid panel have been successfully carried out. The first four natural frequencies and mode shapes were identified. A simple analytical model developed in the study provides good estimates of the natural frequencies of the bending modes. Bending and torsional modes were analyzed with a simple finite element model. The predicted mode shapes agreed very well with those observed experimentally. Failure loads in the compression test were lower than predicted. This may have been caused by problems with the fixtures used to hold the specimen. Predicted strains were in good agreement with the average values measured in the test. The strains varied more with location on the panel than anticipated in the model, resulting in some strains being higher and others lower than predicted.

SOLAR SCINTILLATION AND THE MONITORING
OF SOLAR SEEING

Edward J. Seykora
Professor
Department of Physics

East Carolina University
Greenville, NC 27834

Final Report for:
Summer Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D. C.

July, 1992

Solar Scintillation and the Monitoring
of Solar Seeing

Edward J. Seykora
Professor
Department of Physics
East Carolina University
Greenville, NC 27834

Abstract

A non-telescopic method of determining the quality of atmospheric seeing is discussed for large angular diameter objects, such as the sun. In this method the concept of thermodynamic fluctuations is used to relate the observed intensity fluctuations to the RMS angular diameter of the atmospheric seeing cells and telescopic angle of arrival fluctuations. Comparisons between the measured RMS cell size and telescopic angle of arrival fluctuations are presented for various degrees of seeing. Cross correlation coefficients of 0.95 have been measured during such comparisons.

Solar Scintillation and the Monitoring Of Solar Seeing

Edward J. Seykora

INTRODUCTION

As a consequence of the random motion of air cells with different temperatures and densities, the index of refraction varies from point to point throughout the earth's atmosphere. Such random fluctuations in the index of refraction result in random fluctuations in the direction of light from a distant source, which is referred to as "seeing". In contrast, random intensity fluctuations of the light intensity received are referred to as "scintillation". Because of the importance of seeing causing distortion of ground-based telescopic images, numerous investigations have been carried out concerning quantitative measurements of its magnitude, both for night- and day-time observations (c.f. Roddier, 1981). To date, the methods used in optical investigations of solar seeing utilize telescopic images of the solar disk and depend on: photoelectric measurements of solar granulation contrast (Zindel, 1963); speckle-interferometry of the granulation (Ricort and Aime, 1979); and measurements of fluctuations of the wavefront angle-of-arrival from the solar limb (Brandt, 1987).

Such telescopic measurements of the solar seeing, although accurate, suffer from several drawbacks, such as:

1. The requirement of a telescope, which may be inconvenient in a preliminary site testing campaign.
2. Telescopic image-plane measurements are typically very sensitive to guiding errors and vibrations.
3. The contribution of the telescope itself to the seeing measurement through convection in the telescope tube, and thermal differentials between the telescope structure and surrounding air.

STATISTICAL REPRESENTATION OF SOLAR SCINTILLATIONS

In this investigation, atmospheric seeing and the associated solar scintillations are treated from a statistical point of view such that the small-scale turbulent atmosphere consists of a large number of cells with various sizes, densities, and refractive indices. Here the average cell size (angular diameter) and number will be estimated from the intensity fluctuation measured at the surface of the earth for a large angular diameter source, such as the sun.

Most physical quantities in a thermodynamic system consisting of a large number of particles, or cells in this case, show a Gaussian or normal distribution, and the relative fluctuations of a quantity, G ,

will usually be given by an equation of the form:

$$\frac{\langle G^2 \rangle - \langle G \rangle^2}{\langle G \rangle^2} = \frac{1}{\langle G \rangle}, \quad (1)$$

where the $\langle \rangle$ denotes average values. For most physical systems $\langle G \rangle$ is proportional, or equal, to the number of particles in the system. The density and temperature fluctuations would then be expected to be of the form,

$$\frac{\rho - \langle \rho \rangle}{\langle \rho \rangle} = \frac{T - \langle T \rangle}{\langle T \rangle} = \frac{1}{\sqrt{N}}, \quad (2)$$

where N represents the number of cells in the system. Here it has been assumed the turbulent cell motions are adiabatic in that the cells result from temperature differences and not pressure differences. Under such conditions for an ideal gas, $d\rho/\rho = -dT/T$, which in a statistical sense is consistent with Equation 2. Likewise, it is assumed the intensity fluctuations or scintillations of the sun, as measured at a point in the atmosphere, will follow the equation

$$\frac{I - \langle I \rangle}{\langle I \rangle} = \frac{\Delta I}{\langle I \rangle} = \frac{1}{\sqrt{N}}, \quad (3)$$

where I represents the measured intensity and N the number of seeing cells that cover the solar disk. The angular area of an average (RMS) cell is then,

$$A = \frac{\pi \left(\frac{\theta}{2}\right)^2}{N} = \frac{\pi \theta^2 \left(\frac{\Delta I}{\langle I \rangle}\right)^2}{4}, \quad (4)$$

where θ is the angular diameter of the solar disk. For a long time average of all cells, it is assumed that the average cell is circular, with angular area $\frac{\pi \omega^2}{4}$; then the angular diameter of the average seeing cell is,

$$\omega = \theta \left[\frac{\Delta I}{\langle I \rangle} \right]. \quad (5)$$

Here, ω , can also be considered as the spread or blurring angle of features on the solar disk, or the disk itself.

Several points are worth noting in relation to the above approach to solar scintillations. Under conditions of good seeing, $\omega = 1$ arcsecond, $\Delta I/\langle I \rangle = 5.3 \times 10^{-4}$. From Equation (2), under the same conditions and at an ambient temperature of 27°C , the temperature variations of the seeing cells are of the order of 0.01°C . This value for temperature variations is in agreement with that predicted from

starlight scintillations using a different formalism (Reiger 1963). Likewise, the refractive index fluctuations are also in agreement, in that for an adiabatic process, the index of refraction fluctuations are directly dependent on the average air density and $T/\langle T \rangle$.

It should also be noted that in the above example, where $\omega = 1$ arcsecond, there are 3.6×10^6 cells which silhouette the sun from the point of view of the detector. In order for Equation (1) to be statistically valid, N must be large, which is the case for large angular diameter sources, such as the sun or possibly the moon, and typical seeing situations, ranging from a fraction of an arcsecond to multiple arcseconds. For small angular diameter sources, stars or planets, Equation (1) is invalid in that N , for typical seeing conditions, is a small number. From Equation (5) it is apparent that for best seeing, $\Delta I/\langle I \rangle$ is small, which implies a large N . An interesting limit is the case where N is very large. In this limit with no small-scale turbulent cells, that is, an isothermal atmosphere, the ever present density fluctuations of the atmospheric gas produce density or volume fluctuations having linear dimensions of the order of a wavelength of light. The net result of which leads to variations of the index of refraction throughout the atmosphere in sufficient amount to cause scattering of light in a mathematically equivalent form to that of Rayleigh scattering (c.f. Sears). Evidently, the density fluctuations that persist throughout the atmosphere, range from cells with dimensions on the order of the wavelength of light (which produce scattering), to cells with dimensions of centimeters or meters, which affect the atmospheric seeing.

Up to this point, only a small, or point detector has been considered for measurements of the RMS seeing cell angular diameter in terms of the intensity fluctuations. For larger apertures, it would be expected that fluctuations of the intensity would be reduced. Scintillation light patches at the detector plane, which are smaller than the diameter of the detector, would average, such that the integrated intensity fluctuations approach zero. The predominate intensity fluctuation would then result from scintillation patches with diameters larger than the detector diameter. For atmospheric cells at a height, h , above the detector, the diameter of the light intensity disturbance at the detector is $d = h\theta$. Here any effects resulting from diffraction are considered to be small compared to θ for heights of typical interest, since the angular spread is of order λ/d .

It would appear, because of the spatial filtering imposed by a large aperture detector, that the seeing may be measured or estimated for various levels in the atmosphere from ground level. Such measurements would allow performance comparisons of seeing between ground-based and tower telescopes.

INSTRUMENTATION

The small detector assembly used for measurements of the fractional deviation of the solar scintillations consisted of a Si photovoltaic detection cell filtered by a combination of Schott filters, BG-18 and KG-3. This combination of filters resulted in a spectral passband of 220 nm half-width

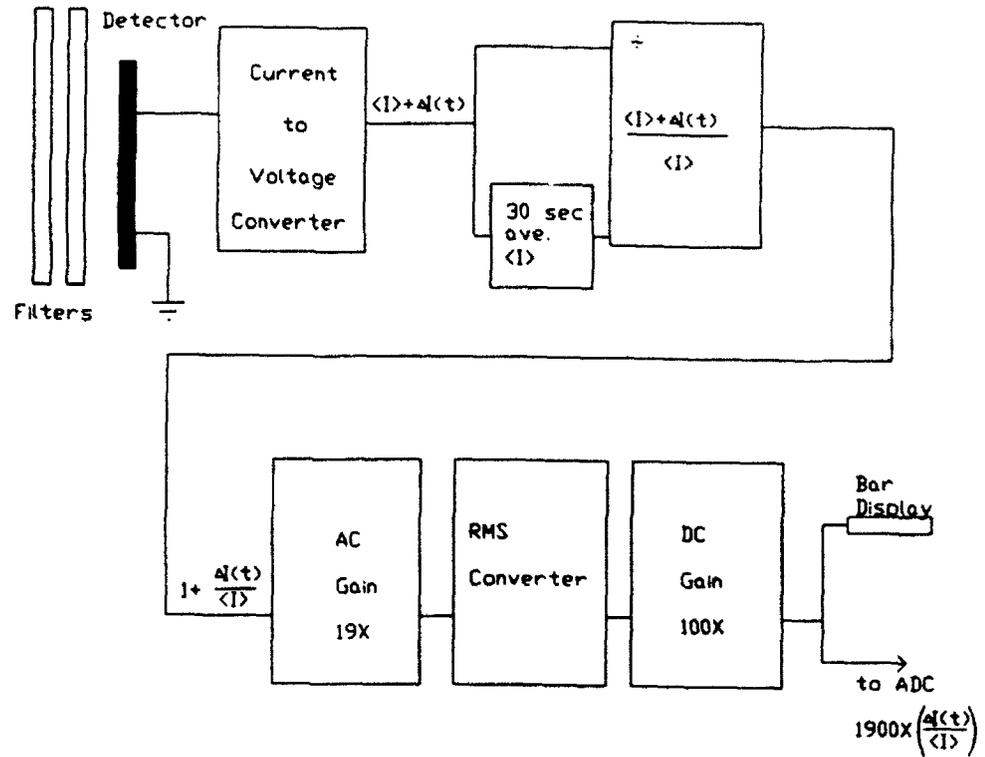


Fig. 1: Schematic representation of circuitry used for measurements of the RMS seeing cell angular diameter and the telescopic RMS angle-of-arrival fluctuations.

centered at 510 nm. For solar tracking, the detector assembly was mounted on the coronagraph spar adjacent to the coronagraph aperture at the Evans Facility. The output current of the detectors was connected to an analog computer circuit for processing as shown in Fig. 1. In the first stage the detector output current, which was proportional to the average light intensity $\langle I \rangle$ plus $\Delta I(t)$ (the fluctuation about the mean), was converted to a voltage signal by a current to voltage converter. This output voltage, $\langle I \rangle + \Delta I(t)$, was then divided by the average (RC time constant 30 sec) voltage, $\langle I \rangle$, yielding an output of $1 \text{ volt} + \Delta I(t)/\langle I \rangle$. After an AC coupled amplification of 19X, the RMS fluctuation, $19X(\Delta I(t)/\langle I \rangle)$, was determined using an RMS converter and followed by a final DC gain stage of 100 times. The net effect was to produce an output voltage, calibrated in volts, given by:

$$V=1900(\Delta I(t)/\langle I \rangle).$$

Thus the output voltage was equal to the RMS seeing angle where 1 volt = 1 arcsecond. The frequency pass band of the amplifier systems, which respond to the intensity variations, was constant from 0.2 Hz to 20 kHz. This insured that the major components of the intensity fluctuations, which contribute to the seeing variations, were recorded. The output voltage representing the seeing was digitized and computer recorded, as well as displayed directly on a LED bar display.

OBSERVATIONS

As mentioned in Section 2, it was assumed that the seeing cells follow a Gaussian distribution. In order for Equation 1 to be statistically valid, the seeing cell distribution function should be equal to, or approximate, a Gaussian distribution provided the cells are uncorrelated. For measurements taken over a short time interval, in comparison to daily atmospheric trends, the correlation between cells would be expected to be small.

Shown in Figure 2 is a typical seeing frequency distribution derived from 300 data points, which were recorded over a one hour period. The mean for the seeing during this period was 0.95 arcseconds with a standard deviation of 0.21 arcseconds. Also plotted in Figure 2 is a Gaussian distribution for comparison with the same mean and standard deviation. As expected, the recorded distributions closely follow the Gaussian distribution.

For comparisons between the RMS angular seeing—cell size and telescopic angle—of—arrival measurements, a slit detection system was used on the primary image of the coronagraph. The slit was arranged perpendicular to the solar limb such that 0.01 m of the solar limb covered the 0.02 m slit. Identical filters and circuitry were used to that shown in Figure 1, with the exception that the amplifier gains were adjusted such that 1 arcsecond was equivalent to 1 volt at the output. This arrangement was very similar in configuration to that used by Brandt, 1987.

Shown in Figure 3 is a scatter diagram of the RMS seeing cell angular diameters vs. the RMS

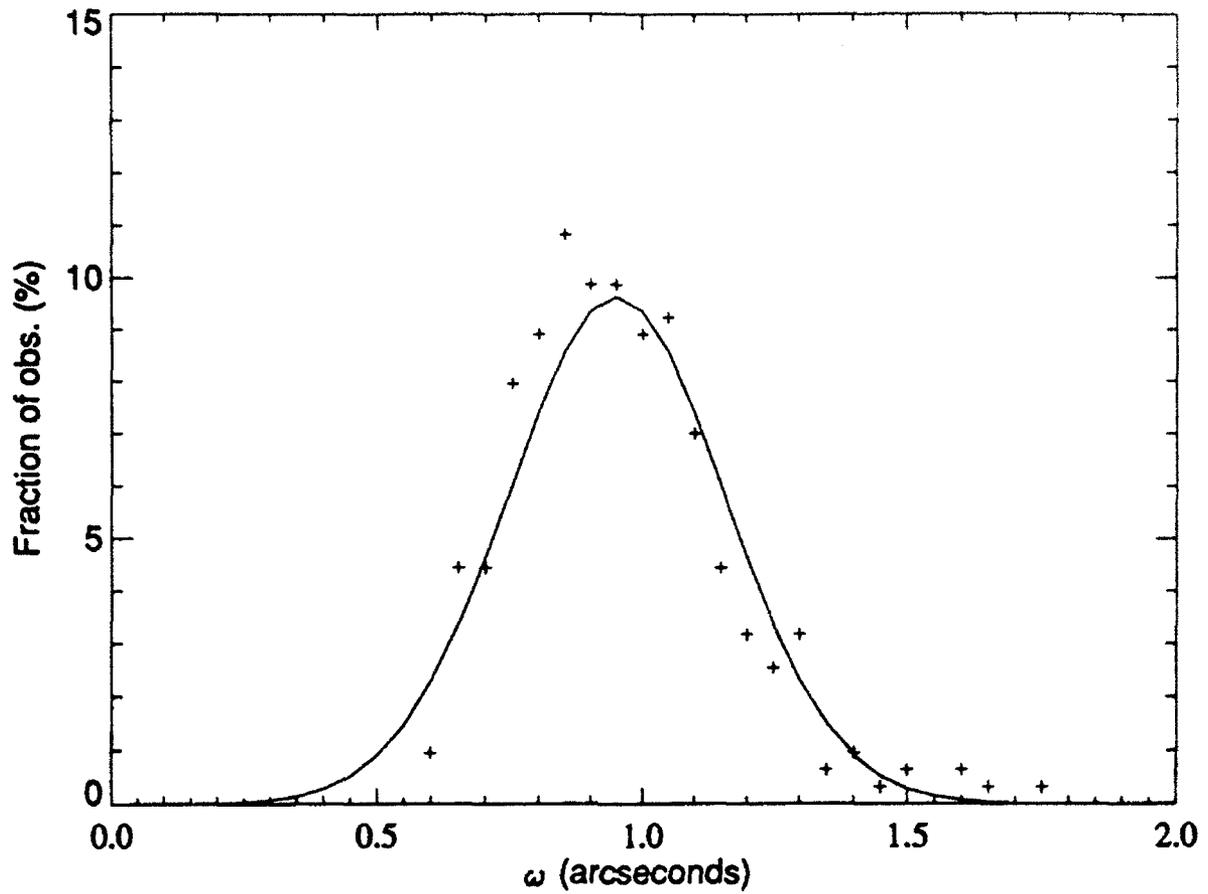


Fig. 2: Frequency distribution of ω , the RMS seeing cell angular diameter in arcseconds vs. fraction of observations + + + +, and a Gaussian distribution with a similar mean and standard deviation ———.

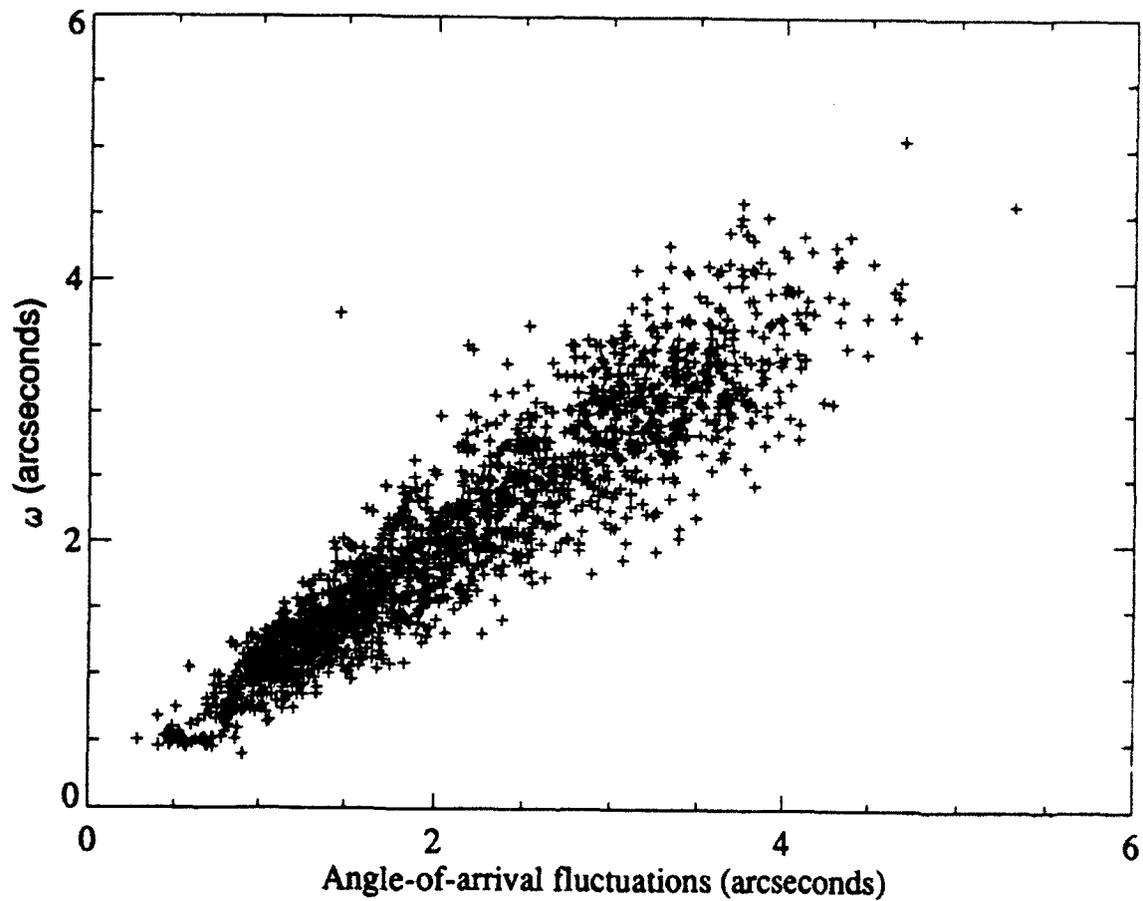


Fig. 3: Scatter plot of the RMS angle-of-arrival fluctuations vs. ω , the RMS seeing cell angular diameter in arcseconds.

angle-of-arrival variations, as measured in arcseconds. This data were recorded under various seeing conditions over a six day period in May and June 1992. Each data point represents the average seeing over a one minute interval. The six data sets used were taken from days that had a large variation in seeing conditions, such that a greater range of seeing values were available for cross correlation. For the 2036 data points shown in Figure 3, a cross correlation coefficient of 0.95 was calculated, indicating a very high degree of correlation, and a high probability that each detector system was measuring the same quantity. Here it is noted that the angle-of-arrival method for seeing measurements results from a small angular diameter section of the solar disk, whereas scintillation measurements provide an RMS average seeing for the full disk. Both methods should be expected to converge to the same value, provided the RMS average is taken over a sufficiently long period in comparison to the longest period of the typical seeing variations.

In order to verify that estimates of the seeing, at a given altitude above ground level, can be carried out through spatial filtering, three scintillation seeing monitors were used. Two detectors were placed at the top of the spar of the Evans Facility. One of these detectors (0.01-m aperture) was used to record the RMS seeing at the coronagraph aperture, as discussed in Section 3.2, and provided seeing measurements at essentially ground level. The second detector with identical filters was arranged such that the reflected light from a white plate, mounted perpendicular to the sun, was imaged with a short focal length lens on the detector. The effective aperture of this detector was set such that a 0.35-m diameter on the reflection surface was imaged on the detector. With this effective aperture of 0.35 m, the recorded scintillations would be expected to have their origins from heights of 40 m or higher. For comparisons, a third detector (0.01 m aperture) was placed at the vacuum window periscope feed of the Vacuum Tower Telescope. This feed provided an unvignetted view of the sun directly from the coelostat mirrors. The horizontal separation of the Evans Facility and the Tower Telescope is 140 m, and the effective height difference between the detectors was 40 meters. Although the detectors at the Evans Facility and the Tower Telescope were separated by 140 m, it may be expected that the projected seeing, as measured with the large-aperture detector, should correlate with seeing at the Tower height. For most of the day each of these detectors would be detecting scintillations and seeing conditions above the ground level atmospheric boundary layer.

The seeing in arcseconds, as recorded by each of the three detectors on 1 July 1992, is shown in Figure 4. Several features are apparent from the time sequences shown in Figure 4. The typical seeing at the Evans Facility in the early morning hours is comparable to the seeing at the Tower Telescope. For later hours, the seeing degrades rather rapidly through midday, generally followed by a modest improvement later in the day. Here the ground or near-ground level microturbulence predominates, as driven by the solar heating of the ground surface and local environment. Above this microturbulent boundary, as expected, the seeing improves. As may be noted in Figure 4., the seeing at the Tower Telescope, and that predicted for the height from the large aperture detector, show greatly improved

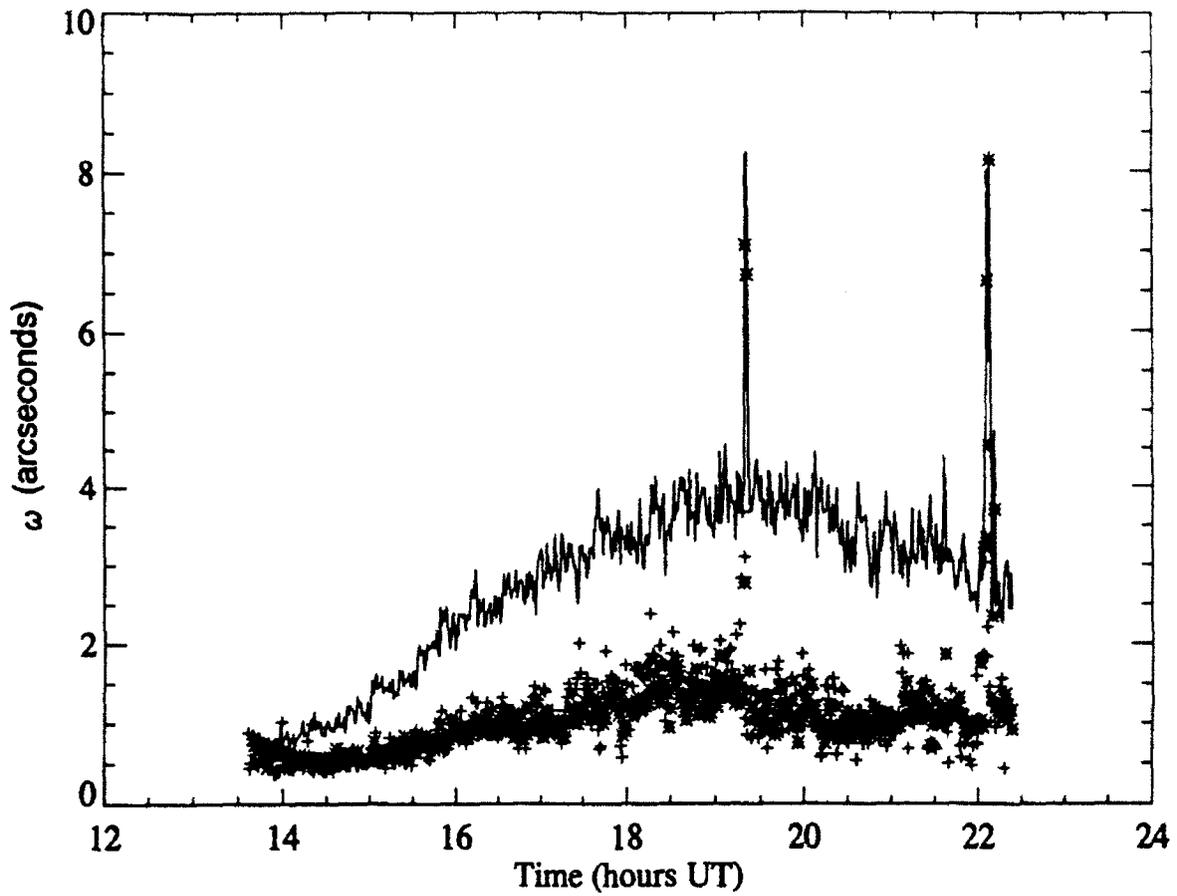


Fig. 4: RMS seeing cell angular diameter vs. time recorded on 1 July 1992: at the Evans Facility ———, Vacuum Tower Telescope + + + +, and predicted for the Vacuum Tower Telescope from the Evans Facility using spatial filtering * * * *.

seeing.

The time sequences of the seeing at the large aperture detector and tower follow one another very closely for this data set, and for seven other days tested. The data set corresponding to Figure 4. had one of the largest variations in the seeing throughout the day and is shown here to allow a clearer representation of the correlation between the predicted and measured seeing at the Tower. The two large spikes in the data resulted from the passage of clouds through the field-of-view of the detectors for several minutes in each case. Nevertheless, the cross correlation between the seeing recorded from the two detectors was always high, and in this case, the correlation coefficient was 0.8.

CONCLUSIONS

In this investigation it has been demonstrated that the statistics of solar scintillations may be used to measure the ground level seeing with a small detector, and that this method has a high correlation with a more traditional telescopic approach. Distinct advantages of the scintillation method are its ease of use, in that a telescope is not required, while the method gives an intrinsic measurement of the atmospheric blurring or seeing angle. The use of large aperture detectors, to spatially filter the ground level scintillation patterns, appears to provide a simple method of estimating the seeing at a point well above ground level.

Although the purpose of this investigation was directed toward measurements of solar seeing, this method may have potential for measurements of vertical turbulence profiles in terms of $C_n^2(h)$, the structure constant of atmospheric refractive-index fluctuations. Such measurements of $C_n^2(h)$ may provide direct measurements of atmospheric turbulence, as introduced by Fried (1966), for telescopic images observed through turbulence. In this investigation the RMS seeing cell angular diameter may be converted to Fried's parameter, r_0 , using the relationship between the angle-of-arrival fluctuations and r_0 , derived by Borgnino et al. (1979), and used by Brandt (1987).

ACKNOWLEDGEMENTS

The author wishes to thank Drs. R. N. Smartt and S. Keil for their helpful discussions. The assistance of L. B. Gilliam, J. Elrod and B. Armstrong at the Evans Facility, along with D. Mann and R. Coulter at Tower Telescope Facility during the observations, was greatly appreciated.

REFERENCES

- Borgnino, J., Aime, C., and Ricort, G.: 1979, *Solar Phys.* **64**, 403.
Brandt, P. N.: 1987, *Astron. Astrophys.* **188**, 163.
Fried, D. L.: 1966, *J. Opt. Soc. Am.* **56**, 1372.
Ricort, G. and Aime, C.: 1979, *Astron. Astrophys.* **76**, 324.

Roddier, F.: 1981, in *Progr. Optics*, Vol. XIX, ed. E. Wolf, p. 238.

Sears, F. W.: 1959, *An Introduction to Thermodynamics, The Kinetic Theory of Gases and Statistical Mechanics*, Addison-Wesley, Reading, M.A.

Zindel, D.: 1963, *Z. Astrophys.* 57, 82.

FINAL REPORT
SUMMER FACULTY RESEARCH PROGRAM
AT
PHILLIPS LABORATORY
KIRTLAND AIR FORCE BASE
ALBUQUERQUE NEW MEXICO

September 30, 1992

SPACE DEBRIS IMPACT EFFECTS ON SPACECRAFT MATERIALS

by

Jon H. Shively
School of Engineering
California State University at Northridge

1.0 SUMMARY

Tests to simulate the effects of hypervelocity collisions of space debris were conducted at Los Alamos National Laboratory. The collisions were produced by propelling a thin plate of aluminum in air and in vacuum at targets of polysulfone and a polysulfone graphite fiber composite. The chemical species in the blow-off produced by the collision were measured using a time-of-flight mass spectrometer. Also the resultant damage was examined using optical and scanning electron microscopes.

In addition to the testing, a review of data on hypervelocity shock led to the development of a correlation of shock data for homogeneous materials in which all materials follow a master curve. The basis of the correlation follows from theoretical predictions in the literature in which shock can be represented using the principle of corresponding states. The ramifications of the correlation mean that temperature, pressure, and density can be represented by a single equation-of-state.

2.0 INTRODUCTION

Space debris arise from many sources; spent motor cases, de-orbiting satellites, exploded satellites, space shuttle discards, are a few examples. The disintegration of major components produces chunks of materials which continue to orbit the earth and which further subdivide due to debris interaction. The size distribution ranges from whole components to parts only a millimeter in diameter. Space debris are not particularly a problem unless they happen to collide with spacecraft. If this occurs the resulting damage can be very serious with the possible loss of mission. Some debris particles are estimated to be travelling at hypervelocities of 3-15 km/sec. It is estimated that the energy in the collision of a piece of aluminum only one centimeter in diameter is equivalent to that of a car travelling sixty miles per hour. No wonder spacecraft designers worry about survivability of a spacecraft facing collision with this energy.

The Summer Faculty Research project which is supported by RDL and the Phillips Laboratory took place over 12 weeks during parts of May, June, July, and August, 1992. The project is aimed at developing an approach to studying damage to spacecraft materials from collisions of small particles of space debris traveling at hypervelocities. Originally, tests using simulated space debris launched by an electric gun were planned. However, as a result of further evaluation of the alternatives to the electric gun, a new way of launching simulated space debris was determined which uses a pulsed laser to launch particles. The laser-launched flyer-plate conveniently interfaces with the vacuum chamber which houses the Time-Of-Flight (TOF) mass spectrometer which is used to evaluate the shock induced chemical species. It also meant that the project would be able to be completed on time without extensive modifications to the Los Alamos equipment. The mass spectrometer and the laser launch apparatus are both located at Los Alamos National Laboratory. As a faculty researcher at Phillips Laboratory I spent the summer planning, designing, coordinating, and evaluating the results of both mass spectrometer experiments and shock damage experiments which were conducted at LANL. In addition, time was spent developing an evaluation of shock physics data for a variety of materials. This evaluation, combined with a thorough review of shock physics papers, lead to an interesting and novel correlation of shock properties which heretofore

was unknown. A description of the tasks performed this summer and the results are summarized below.

3.0 TASK 1-THE CHEMISTRY OF SHOCK EJECTA

About 7 years ago, NASA launched an earth satellite called the Long Duration Exposure Facility, LDEF. The Phillips Laboratory participates in that program and is evaluating specimens which experienced a long term exposure to the space environment. The materials includes metals and composites. The Phillips Laboratory specimens were exposed to the trailing and leading edges of LDEF. Post-mortem studies of the specimens revealed several salient features, craters from the impact of small space objects, erosion on surfaces, and surface deposits. The presence of the deposits on originally clean surfaces suggests that chemical reactions are taking place in outer space on surfaces as the result of the migration of chemical species around the satellite. As a result, an experiment was conceived to study the production of chemical compounds from impacts of space debris.

The LDEF specimens of polysulfone were of particular interest because the crystalline deposits found on the aluminum surface contained sulphur. The source of the sulphur is unknown. One possible explanation is that the sulphur came from a specimen of polysulfone composite next to the aluminum specimen upon which were found the deposits. To further speculate, the sulphur could have come from the de-polymerization of the composite resin as a result of hypervelocity impact. Therefore, an experiment was conceived to measure the chemical species which form during shock impact. The species would be detected in the ejecta of the collision process. The technique for accomplishing these measurements is described below.

The experimental setup to measure the chemical species which form after hypervelocity collision with simulated space debris is shown in Figure 1. A YAG laser beam is pulsed for 0.2 ms on a region 600 microns in diameter on the back side of an aluminum-coated quartz substrate. The laser pulse vaporizes the material at the quartz-aluminum interface which causes an almost instantaneous rise in pressure as the plasma forms. The aluminum coating bulges out, breaks away from the substrate, and accelerates very rapidly to terminal velocities between 2 km/s to 10 km/s. The

aluminum flyer then travels a short distance in the vacuum chamber and strikes the target material, polysulfone resin. See Figure 2. The collision produces shock damage in the material including almost instantaneous heating in the polysulfone as result of the shock pressure. The pressure and heating vaporizes the polysulfone which in turn produces chemical changes. Since the vapor is ejected from the ensuing crater, the vapor molecules and ions are thermally quenched as they travel into the throat of the TOF mass spectrometer. The quench stops further chemical reactions from occurring. Chemical species are then ionized and accelerated to the detector of the TOF spectrometer. A typical spectrum is presented in Figure 3.

The spectra of mass 1 to 200 are evaluated every 12 milliseconds so that spectra can be monitored as a function of time. The resultant experimental spectra from the hypervelocity collisions of aluminum with polysulfone showed little change in the distribution of species with time. However, the magnitude of the spectral peaks did increase and remain constant over a few spectral scans. This implies that the species detected formed during the shock event. Therefore, to improve the sensitivity, the spectra were time averaged over several spectral scans of 120 ms.

The time-averaged spectra exhibit some salient features. Clearly the species found were the direct result of the de-polymerization of polysulfone. The presence of radicals associated with the polysulfone mer are evident, i.e., C_2H_2 , aromatic rings, etc. In addition, because the aluminum contains a thin layer of oxide, Al_2O_3 , the subspecies of aluminum oxide also appear in the spectra; i.e., Al_2O , as well as pure aluminum from the flyer plate.

The measurement of the spectra of ejecta from hypervelocity collision by simulated space debris has never been achieved before primarily because conventional methods of producing particles travelling at hypervelocities are not compatible with the use of a TOF mass spectrometer and its vacuum chamber.

4.0 TASK 2-SHOCK DAMAGE IN A POLYSULFONE COMPOSITE

The use of a laser-launched piece of aluminum flyer plate enabled damage to be studied independently of the TOF mass spectrometer and its vacuum chamber. A few collisions were performed at Los Alamos in air on samples of polysulfone resin and a

polysulfone composite. The damage was characterized at Phillips Laboratory using optical and scanning electron microscopy. In comparing the damage to resin with the composite, differences in the resultant morphologies are striking, as discussed below.

The resin craters caused by the hypervelocity collisions were photographed. Typical results are shown in Figure 4. The crater is surrounded by a ring of aluminum fragments. The aluminum flyer plate apparently disintegrates shortly after the collision. The surface of the polysulfone is coated with tiny fragments of the aluminum, especially around the crater. The diameter of the crater is around 650 microns which is very close to the size of the flyer plate. This means that the flyer plate remains intact prior to collision. This has been verified by streak camera pictures taken at Los Alamos.

The craters in the resin were also photographed using the AMRAY scanning electron microscope. A typical crater is shown in Figure 5. The photograph suggest that hypervelocity collisions produce melting in the center of the crater under the region which the disc of aluminum strikes the polysulfone resin. Upon cooling, a crack develops around the periphery of the crater as the neighboring unheated material resists contraction. The edges of the crater are pushed up forming a lip as has been observed in collisions retrieved from space such as on LDEF. The material in the center is also rippled from melting and buckling under the heating process.

The damage observed in the composites differ dramatically from the resin. A typical crater is shown in Figure 6. The crater is characterized by a spalling of the front surface resin and fiber fracturing. The shock damage results in little or no melting even at comparable velocities. This suggests that the shock energy is absorbed by fracture rather than melting and vaporization.

5.0 TASK 3-REVIEW OF THE PHYSICS OF HYPERVELOCITY IMPACT

A significant portion of my research time was spent in a careful review of the literature on the physics of hypervelocity impacts such as those produced by space debris. A major portion of the research in this area was carried out during the 1960's and 70's. About 50 references were reviewed.

In investigating the literature and comparing materials, a interesting discovery was made. It was found that all the data from hypervelocity impacts could be placed on

a single master curve of shock pressure versus particle velocity. The results for a ten homogeneous materials are presented in Figure 7. It should be noted that the pressure is normalized by dividing by the a factor equal to the adiabatic bulk modulus and the particle velocity is normalized by dividing by the bulk sound velocity:

$$B_s = \rho C^2$$

where

B_s	=	Adiabatic Bulk Modulus
ρ	=	Density
C	=	Bulk Sound Velocity

or the Normalized Pressure is the pressure divided by B_s and the normalized particle velocity is normalized by dividing by C .

The reason this correlation appears to work for a large variety of homogeneous materials is due to the hydrodynamic behavior of shock phenomena. As a result, the method of corresponding states can be used to relate thermodynamic variables such a volume, pressure, temperature, and shock velocity. According to this theory all materials when expressed in corresponding states would behave the same. Thus shock thermodynamics for materials would lead to a correlation of shock data which is independent of differences in density, strength, atomic weight and all shock data could be made to fit one graph when the data are normalized.

The master chart can be rationalized as follows. The shock pressure is related to the particle velocity and the shock velocity by the equation

$$P = \rho U_s U_p$$

where

P	=	Pressure of the Shock
ρ	=	Density
U_s	=	Shock Velocity
U_p	=	Particle Velocity.

The shock velocity is apparently related to the particle velocity by an equation of the form

$$U_s = C + sU_p$$

where

s	=	Slope.
-----	---	--------

These equations can be manipulated algebraically to produce a new equation

$$\begin{aligned}
 P &= \rho C^2 (U_p/C) \{1 + s(U_p/C)\} \\
 \text{or } P/\rho C^2 &= (U_p/C) \{1 + s(U_p/C)\}
 \end{aligned}$$

The left hand side of this equation is the reduced pressure and the right hand side is a function of U_p/C which corresponds to the variables in the master plot, Figure 7. Other materials including water have been tested against the correlation and found to follow the same master curve. See Figures 8 to 10. Note that water is included in the list of materials. Since all materials so far have followed the correlation, the value of s must be the same for all materials. To test this hypothesis, a series of graphs of the shock velocity versus the particle velocity were made for several materials. The results are shown in Figure 11. As can be seen, while there is considerable scatter in the data, the slopes of the lines correspond quite well. In fact, comparing aluminum and water we see that the slopes agree quite well with each other and with the theoretical value of 1.33. Thus the behavior under shock is such that materials behave like a highly compressed liquid. Hence, a hydrodynamic analog is appropriate for these materials.

The master chart enables data to be extrapolated from other materials and add confidence to extrapolating data. Since the shock Hugoniots are a vital part of survivability calculations from spacecraft, the master chart reduces errors in estimating shock pressures and temperature.

While damage was carefully characterized during the impact process, the relationship between shock thermodynamics and shock damage is still incomplete and must be the subject of shock experiments in the future.

6.0 CONCLUSIONS

The results of this work, which directly impact Phillips Laboratory's mission of understanding the nature of and protecting against the damage from natural and hostile space environments, are extremely significant. These experimental procedures determine the chemistry of the blow-off and the shock pressure resulting from hypervelocity impacts on spacecraft materials. When combined with Phillips Laboratory's other research efforts of exposing materials to atomic oxygen, ultraviolet radiation and energetic electrons and protons, this laboratory becomes the only facility in this country

able to simulate all important space environmental variables, simultaneously. This major advance in experimental space simulation techniques will allow Phillips Laboratory to develop countermeasures and hardening technologies for space craft materials more accurately and at a much faster pace.

7.0 FUTURE PLANS

The future plans include expanding this research in the area of shock chemistry at Los Alamos and at Phillips Laboratories in the next two years. While reasons for continuing the studies differ at the two laboratories, the cooperative effort demonstrates the extent benefits can accrued both DOD and DOE from this collaboration.

Shock wave chemistry is new and little is known because the techniques for studying it were not available until Phillips Laboratory began this work. Future plans also include a continuation of research at California State University at Northridge, (CSUN). Two areas of collaboration with LANL and Phillips have been identified. CSUN will prepare other specimens for shock studies at LANL and will combine the shock physics with the shock damage studies in order to establish a framework for selecting materials for future spacecraft. It is hoped that the excellent cooperative arrangement between Phillips, LANL, and CSUN can be continued using a Research Initiation Grant as part of this contract.

8.0 RECOMMENDATIONS

The Faculty Summer Research Program funded by AFOSR through RDL is an example of an excellent program. From both the host laboratory and the educational institution perspectives, the faculty research is immensely beneficial. The laboratory gets an opportunity to move into a new research area without diverting valuable personnel and resources away from the main thrust of its mission. The faculty research associate can bring new ideas and expertise to laboratory problems. The educational institution gains by having one of its faculty members exposed to new and innovative technologies without expense and has the opportunity to initiate research on the campus in new areas. The faculty member also gains through experience and has a chance to meet and work closely with other researchers. It is hoped that the AFOSR will continue to

support this very worthwhile program.

It is the authors conclusion that this was a prime example of how the RDL Program should work. It has been very rewarding experience. It is hoped that RDL and Phillips agree. The author acknowledges the support of Dr. Charles Stein at Phillips for his patience and guidance this summer. Also from Phillips the close support of Mr. Robert Roybal and Charles Maglionico are gratefully appreciated. Finally, the recommendations and guidance of Dr. Roy Greiner and Dr. Dennis Paisley at LANL are also appreciated, especially in helping to complete the experiments this summer.

9.0 LIST OF FIGURES

- Figure 1 Experimental Arrangement to Measure Chemical Species in Hypervelocity Shock Ejecta
- Figure 2 Geometrical Arrangement Between the Flyer Plate and the Target After Five Shock Tests in the TOF Cell
- Figure 3 The Spectra of Species Found in the Ejecta of a Hypervelocity Collision of an Aluminum Flyer Plate with Polysulfone Resin
- Figure 4 Optical Photograph of the Crater Created by the Collision of an Aluminum Flyer Plate with Polysulfone Resin
- Figure 5 Scanning Electron Microscope Image of the Crater Created by a Collision of an Aluminum Flyer Plate with Polysulfone Resin
- Figure 6 Scanning Electron Microscope Image of the Damage Created by a Collision of Aluminum Flyer Plate with Polysulfone-Graphite Fiber Composite
- Figure 7 Master Plot of Normalized Pressure for Various Normalized Particle Velocities for Eight Metals
- Figure 8 Master Plot for Several Non Metallic Materials
- Figure 9 Master Plot for Several Plastics
- Figure 10 Master Plot for Water
- Figure 11 Plot of Shock Velocity Versus Particle Velocity for Several Materials

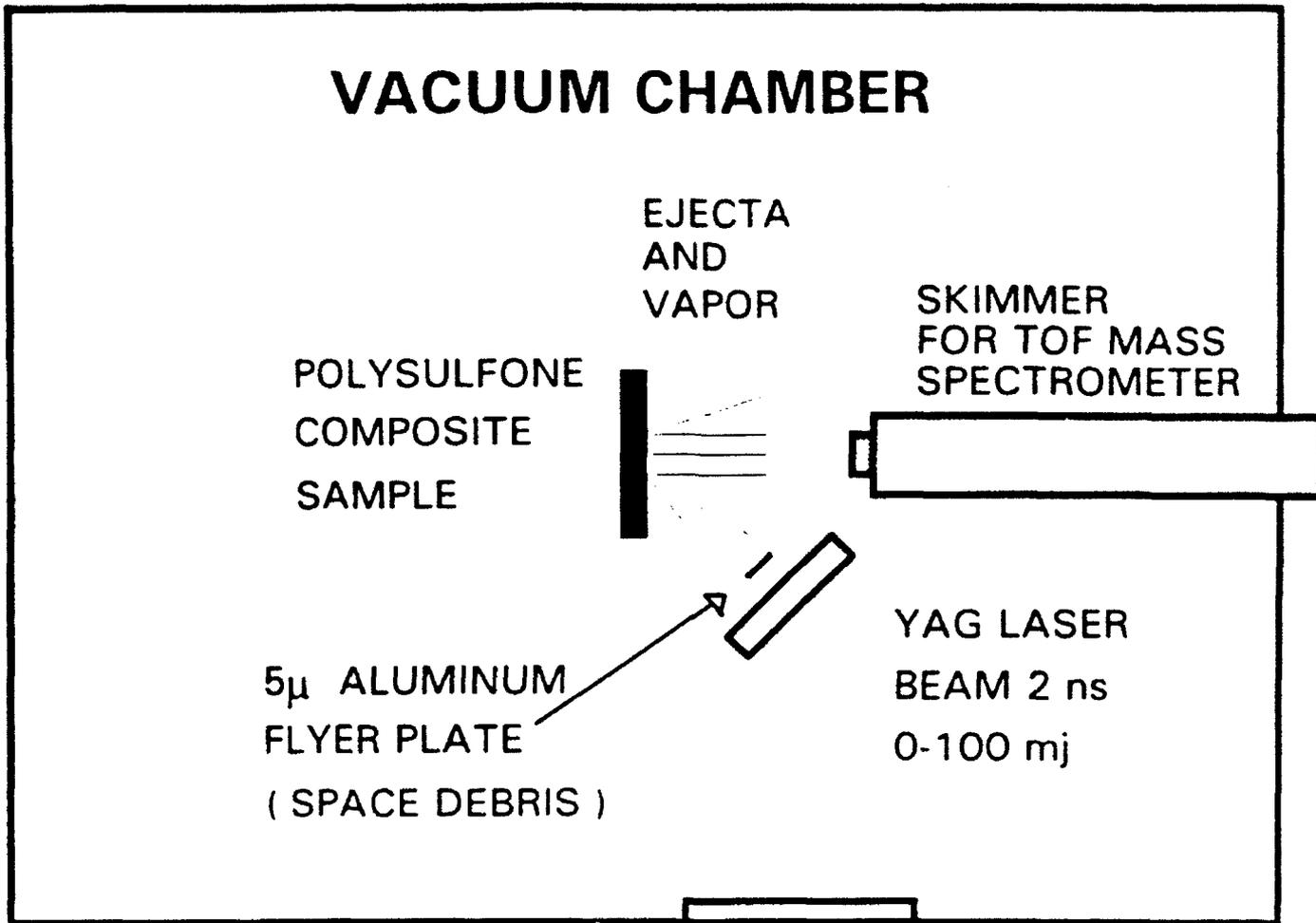
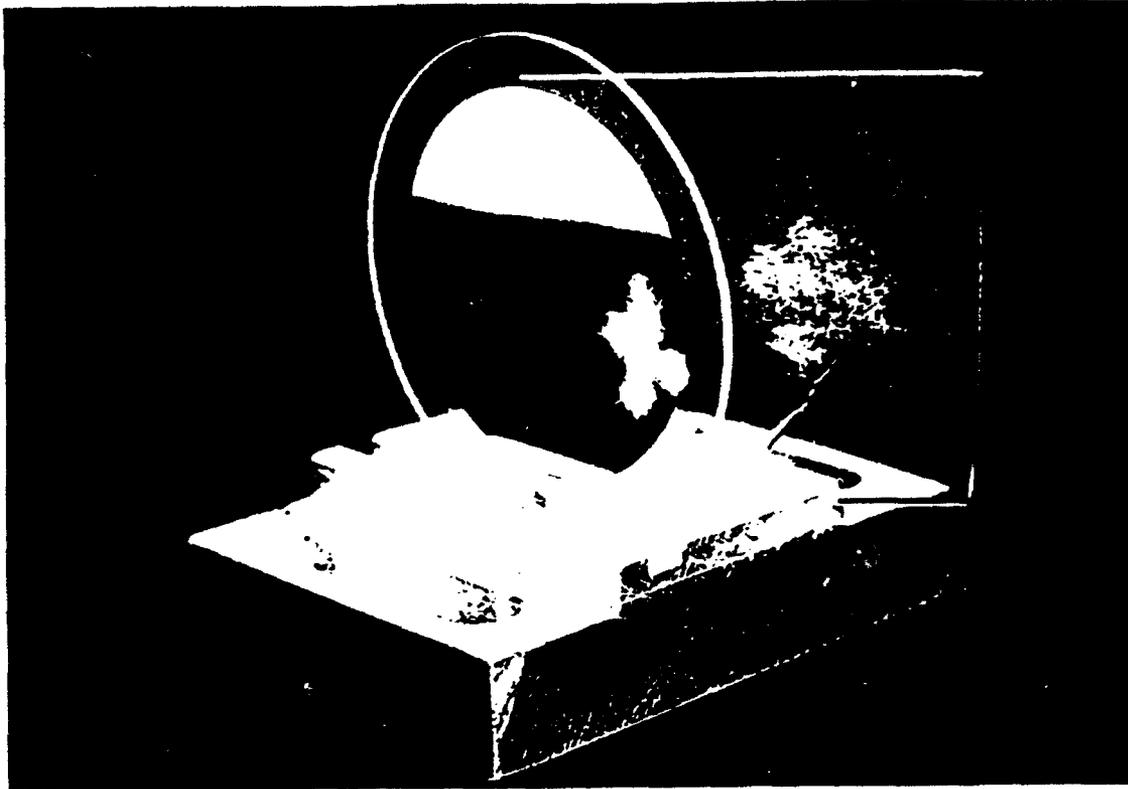


Figure 1 **Experimental Arrangement to Measure Chemical Species in Hypervelocity Shock Ejecta**



**Figure 2 Geometrical Arrangement Between the Flyer Plate and the Target After
Five Shock Tests in the TOF Cell**



Figure 4 **Optical Photograph of the Crater Created by the Collision of an Aluminum Flyer Plate with Polysulfone Resin**



Figure 5 Scanning Electron Microscope Image of the Crater Created by a Collision of an Aluminum Flyer Plate with Polysulfone Resin



Figure 6 Scanning Electron Microscope Image of the Damage Created by a Collision of Aluminum Flyer Plate with Polysulfone-Graphite Fiber Composite

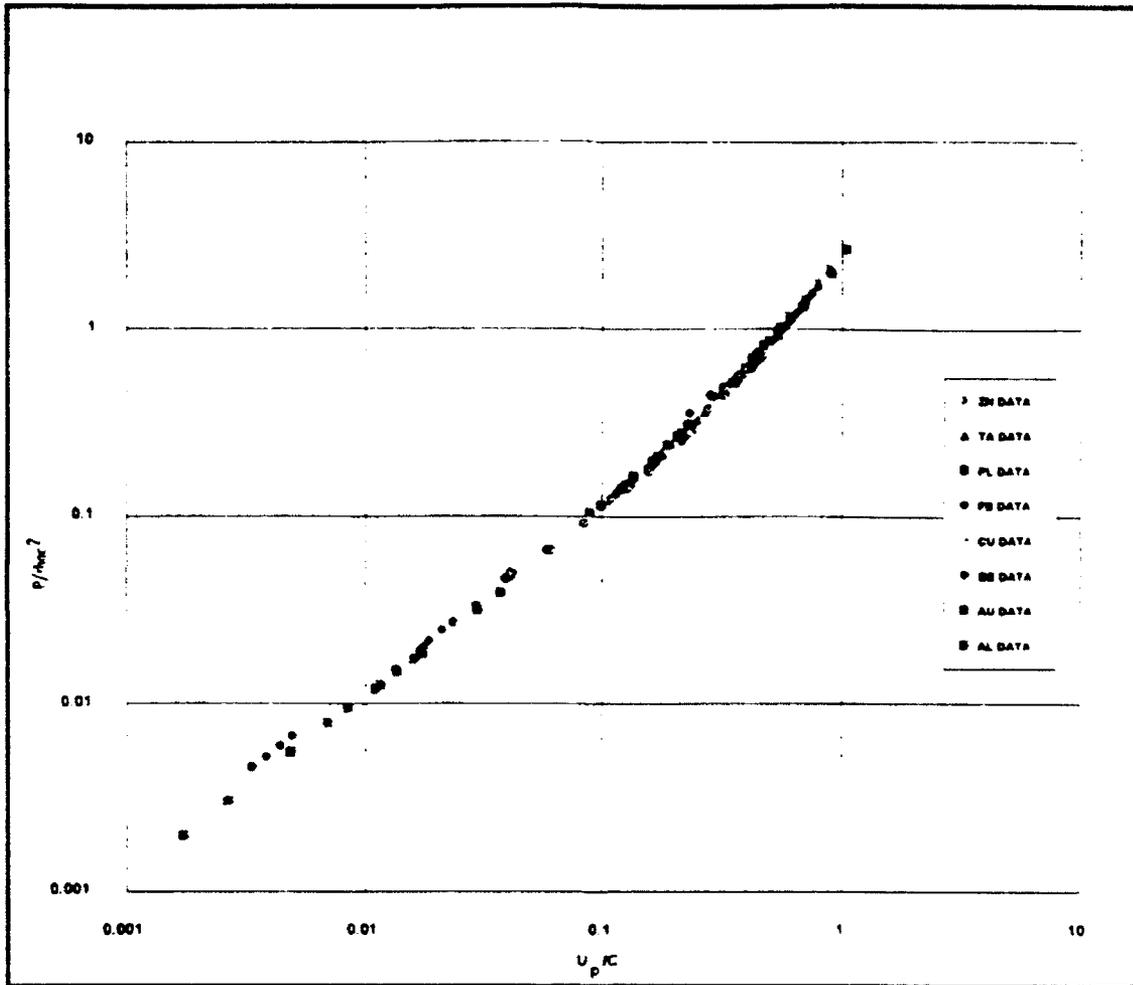


Figure 7 Master Plot of Normalized Pressure for Various Normalized Particle Velocities for Eight Metals

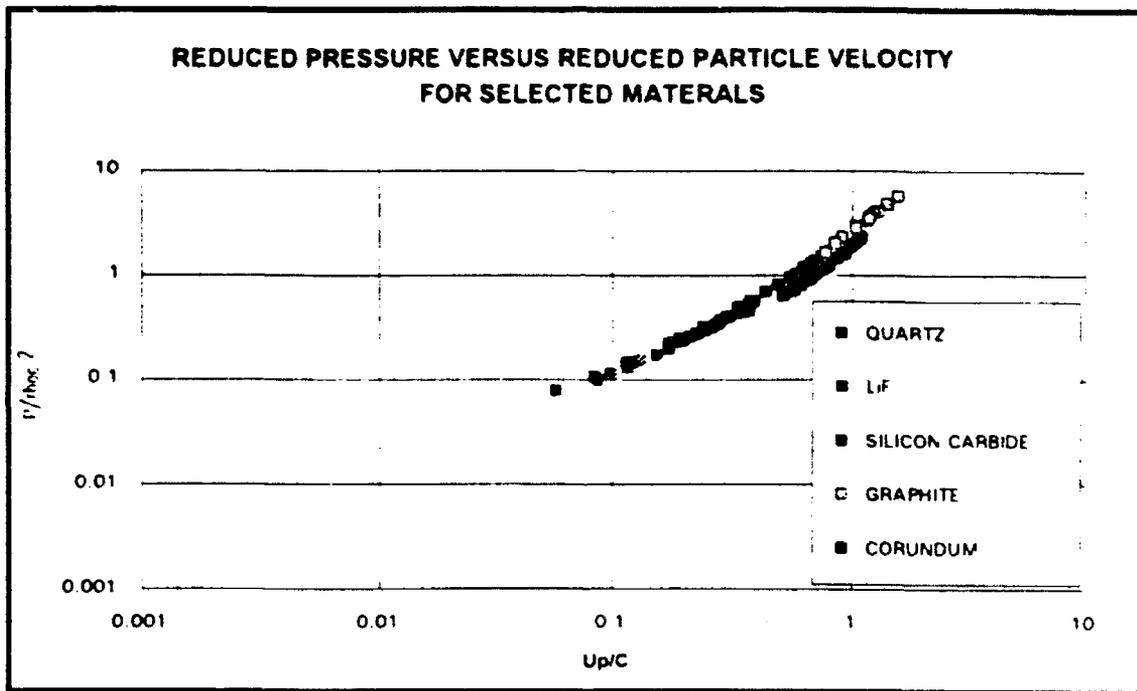


Figure 8 Master Plot for Several Non Metallic Materials

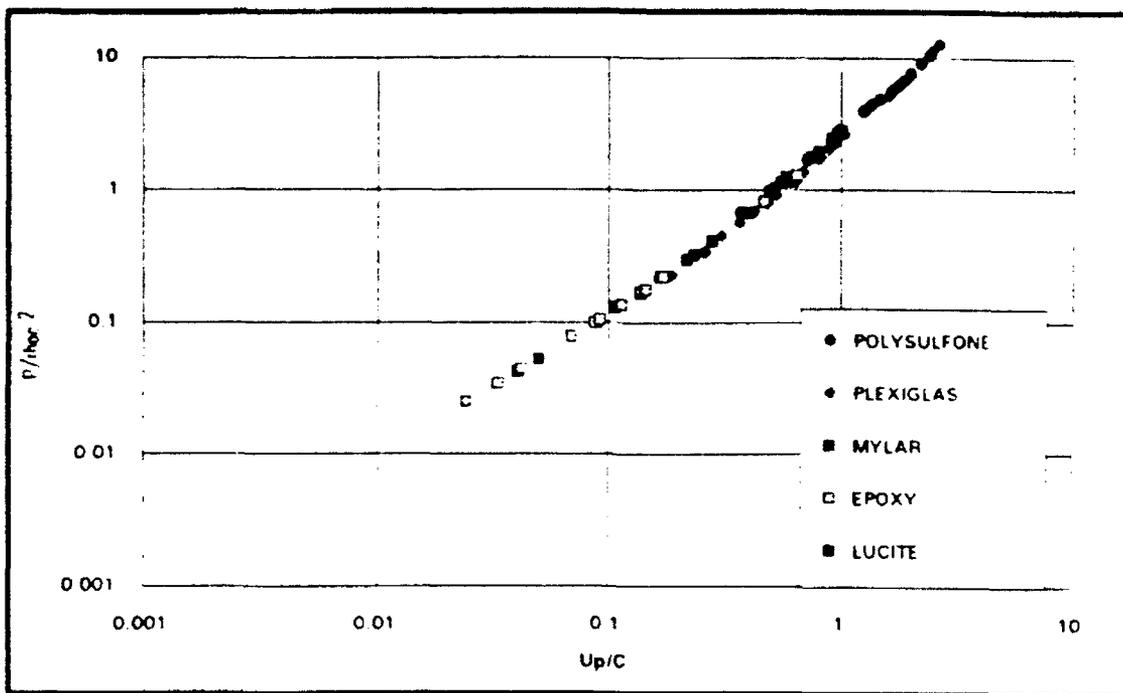


Figure 9 Master Plot for Several Plastics

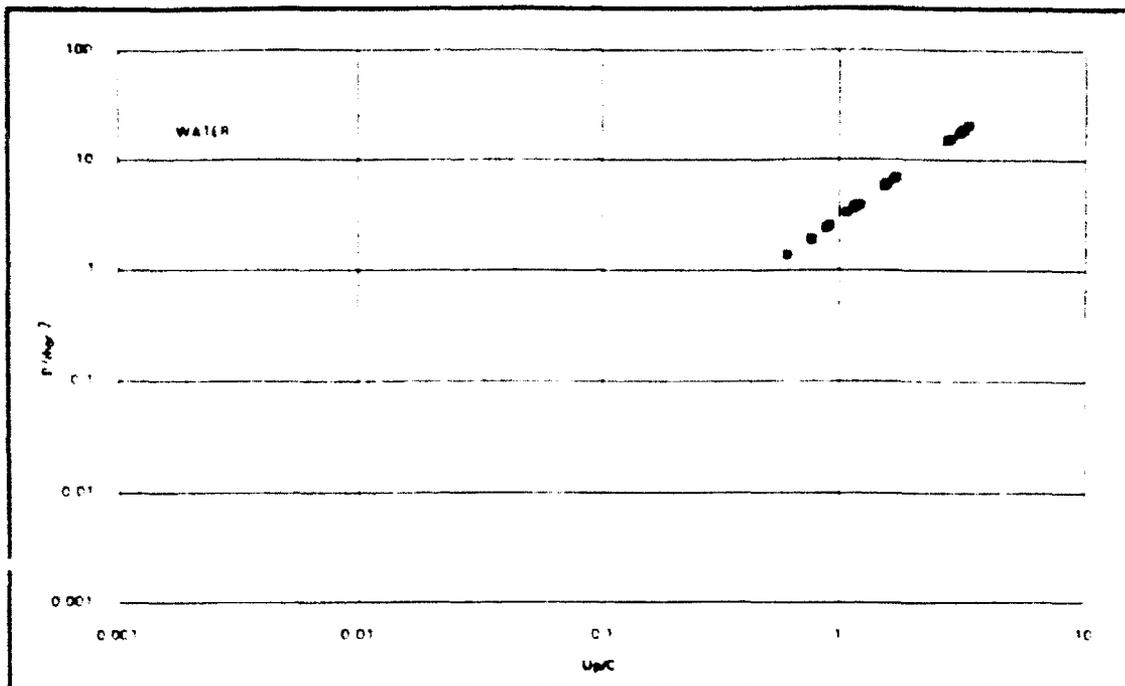


Figure 10 Master Plot for Water

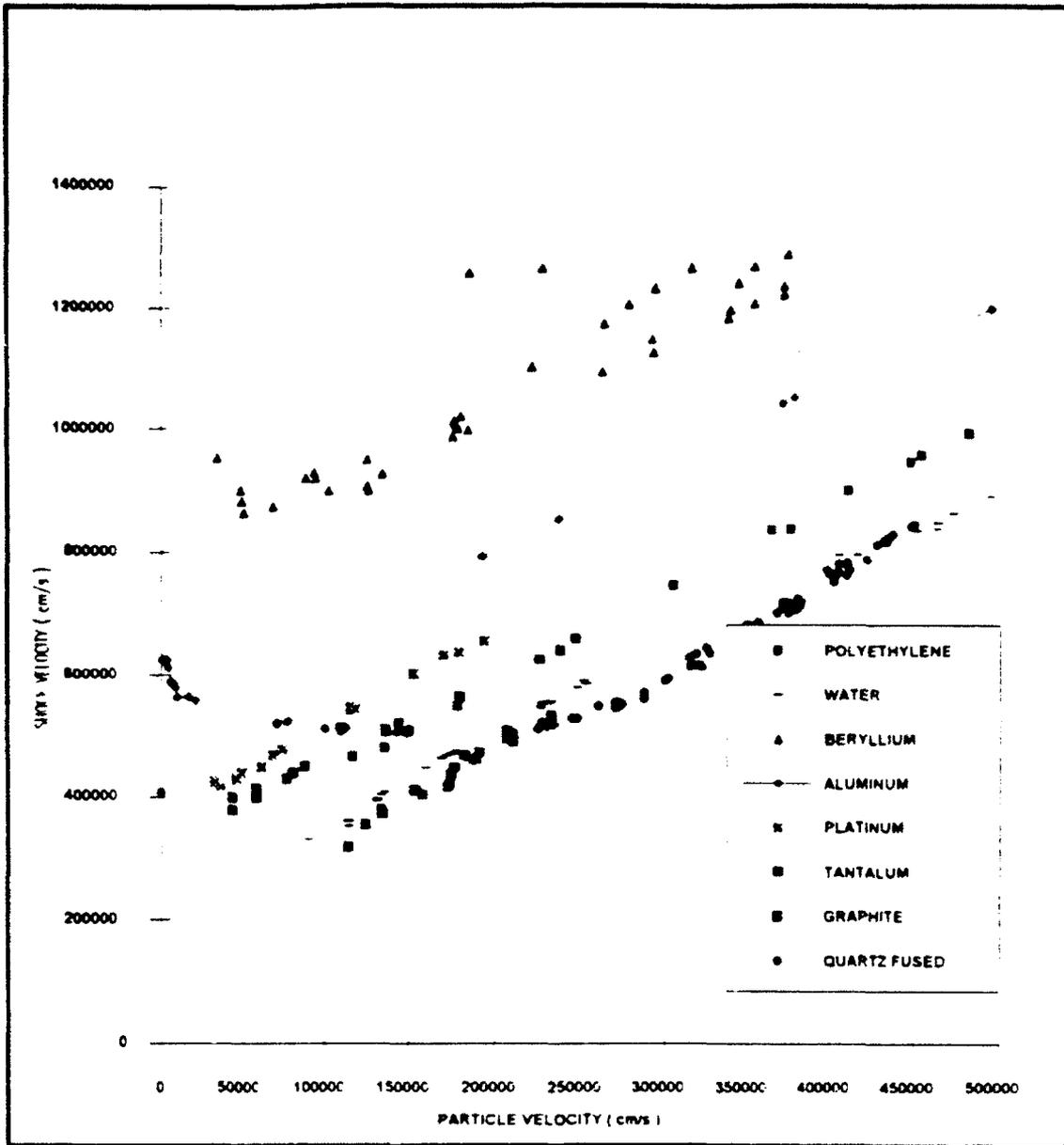


Figure 11 Plot of Shock Velocity Versus Particle Velocity for Several Materials

ESTIMATION OF MISALIGNMENT PARAMETERS
OF MULTI-APERTURE TELESCOPES

Johanna K. H. Stenzel
Associate Professor
Department of Arts and Sciences

University of Houston-Victoria
2506 East Red River
Victoria, Texas 77901

Final Report for:
Summer Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D. C.

September 1992

ESTIMATION OF MISALIGNMENT PARAMETERS
OF MULTI-APERTURE TELESCOPES

Johanna K. H. Stenzel
Associate Professor
Department of Arts and Sciences
University of Houston-Victoria

Abstract

Discrete optimization using the phase-diversity method and optical transfer functions calculated for one hundred piston values in a smaller frequency region yielded very similar piston estimation as did a conjugate-gradient optimization over the entire frequency region. Both methods successfully estimated piston misalignment (within ± 0.1 waves of piston) of an experimental two-aperture telescope for some objects but not others. Wiener-Helstrom filter restoration of the objects showed less noise for the images with piston misalignments that were more successfully estimated.

This paper develops a model for the optical transfer functions of mirrors with arbitrary numbers of apertures and presents a method for reducing the dimension of their misalignment-parameter space and the size of their frequency domain.

ESTIMATION OF MISALIGNMENT PARAMETERS
OF MULTI-APERTURE TELESCOPES

Johanna K. H. Stenzel

I. INTRODUCTION

Phase diversity is a method of estimating misalignment parameters in multi-aperture telescopes developed by Gonsalves [1] and others [2,3,4]. These authors use the method of conjugate-gradients to optimize the misalignment parameters over the frequency domain of the telescope aperture system. The method uses two (or more) images-- one with a deliberate defocusing added to the already present aberrations for phase diversity. For adaptive applications, this method may not be fast enough but perhaps a neural network approach might be. Angel et al. [5] used a supervised perceptron neural net on the two diversity intensity images to estimate the misalignment parameters of a six-mirror modeled telescope. Miller [6] suggested that the method could be made faster by using a neural network on the single ratio of the Fourier transforms of the diversity images instead of on the two images themselves. This ratio is defined in the frequency domain of the mirror aperture but only the subdomain in which the ratio varies with the misalignment parameters need be used. Since this ratio can be measured and is a function of the unknown misalignment parameters, the problem of estimating these parameters can be thought of as an inverse function problem.

Miller and Ling [7] obtained experimental diversity images for a number of objects with a two mirror telescope with preset piston misalignments. Then they estimated the piston misalignments by optimizing (with conjugate-gradients) a variation of the Gonsalves objective function called the Wiener-Helstrom function. This study uses their data from three objects--the first a point source and the other two different views of an Helium-Neon laser illuminated slide of a bar chart. A previously developed model for a two mirror telescope [8] yields ratios of the Fourier transforms of the diversified images of a point source for each value of piston from

-0.49 to 0.50 in steps of 0.01. The next section discusses how a neural network-like table look-up of the computed ratios can estimate the piston parameters from the experimental data. The third section reconstructs the objects with a Wiener-Helstrom filter and estimates the noise of the data from the Fourier transform of the reconstructed objects. The fourth section extends the model of the two mirror telescope [8] to an arbitrary number of symmetrically (circularly placed) circular mirrors. Section 5 then outlines how the parameter dimension and frequency domain may be reduced for optimization of the misalignment parameters.

II. ESTIMATING PISTON MISALIGNMENT FROM EXPERIMENTAL DATA

The experimental telescope had an effective focal length of 2341 mm. and used an aperture mask with aperture radii of 9.1695 mm. and aperture center separations of 21.768 mm. The aperture mask is illustrated by the solid circles 1 and 2 in Figure 1.

The data consists of pairs of image intensity arrays with a diversity of 0.51 waves of defocus and various piston misalignments. For imaging systems with spatially incoherent, quasi-monochromatic light, an object intensity $f(x)$ is related to image intensity $g(x)$ by means of convolution \star with the system's point-spread function $s(x)$ (the image of a point source):

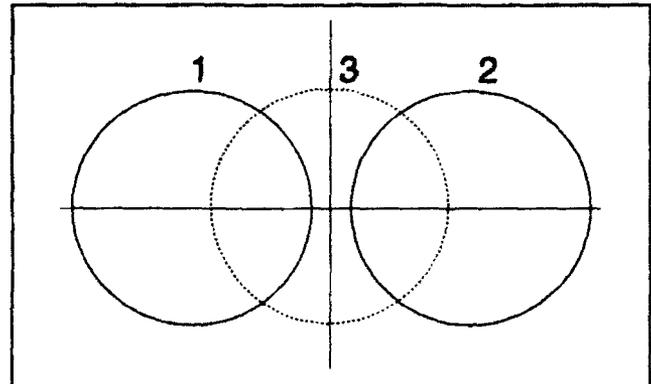


Figure 1. Circles 1 and 2 show the telescope aperture with centers separation 21.768 mm and radii 9.1695 mm. Division by 632.8 nm wavelength and 2341 mm focal length scale it to the frequency plane.

$$g(x) = f(x) \star s(x), \quad (1)$$

in the notation of Paxman and Fienup [2]. The Fourier transform of Eq. (1) to the frequency plane ξ is

$$G(\xi) = F(\xi)S(\xi), \quad (2)$$

where capitalization indicates Fourier transformation and $S(\xi)$ is the optical transfer function (OTF) of the system. If Eq. (2) is evaluated twice for a pair of diversity images, the transform of the object $F(\xi)$ would drop out from the ratio of these, since the object is independent of the optical system and its diversity. Thus, the ratio of the diversity images of any object is equal to the ratio of the OTF's of the system with its diversity. The basis of the phase-diversity method of determining the misalignment parameters is to compare OTF's calculated from a system model to the Fourier transforms of diversity image data. Since the telescope is not identical to its model and the image data includes measurement errors, Paxman, Schulz, and Fienup [4] use the notation d and D for the detected images and their Fourier transforms, respectively:

$$D(\xi) = G(\xi) + N(\xi), \quad (3)$$

where $N(\xi)$ is the Fourier transform of additive noise $n(x)$. Thus, although the ratio of the transforms of the diversity data cannot be expected to equal the ratio of the modeled OTF's, they should be close in some sense.

The usual metric of comparison is the Euclidean norm

$$\sum_{\xi \in \Xi} \left| \frac{D(\xi)}{D_d(\xi)} - \frac{S(\xi)}{S_d(\xi)} \right|^2, \quad (4)$$

where Ξ is the frequency domain and the subscript d refers to diversity defocusing of the image. Zeros in the denominators do not occur if the domain is restricted to the region where the OTF ratio shows any variation with piston (the regions where circles 1 and 3 and 3 and 2 overlap in Figure 1). Otherwise, the denominators would have to be multiplied out. If the additive noise is Gaussian, Gonsalves [9] found a closed form expression for an object that maximizes a log-likelihood function. This leads to a Gonsalves norm [4]

$$\sum_{\xi \in \mathbb{R}^2} \frac{\left| \frac{D(\xi)}{D_d(\xi)} - \frac{S(\xi)}{S_d(\xi)} \right|^2}{|D_d(\xi)|^2 \left(\left| \frac{S(\xi)}{S_d(\xi)} \right|^2 + 1 \right)} \quad (5)$$

A third possible norm arises from the Wiener-Helstrom filter restoration of the Fourier transform of the object as given in [3]

$$F(\xi) = \frac{S^*(\xi)D(\xi) + S_d^*(\xi)D_d(\xi)}{|S(\xi)|^2 + |S_d(\xi)|^2 + \frac{P_n(\xi)}{P_r(\xi)}} \quad (6)$$

where * indicates complex conjugation and P_n and P_r are noise and object power spectral densities, respectively. Paxman and Fienup [2] claim that, in practice, P_n is taken as a constant and P_r is estimated. With P_n/P_r estimated to be

$$\frac{P_n(\xi)}{P_r(\xi)} = \sigma [S(0,0)^2 + S_d(0,0)^2], \quad (7)$$

with σ a smoothing parameter, Ling uses a Wiener-Helstrom norm

$$\sum_{\xi \in \mathbb{R}^2} \frac{|D(\xi)S_d(\xi) - D_d(\xi)S(\xi)|^2}{|S(\xi)|^2 + |S_d(\xi)|^2 + \sigma [S(0,0)^2 + S_d(0,0)^2]} \quad (8)$$

in her conjugate-gradient optimization to estimate the piston parameters of the data with two values of σ : $\sigma=0.00005$ and $\sigma=0.5$. Notice that for this norm, it is not sufficient to use the ratio of the OTF's; Both OTF's, S , with and S without defocussing, are required.

The neural network approach (without training) to estimating the piston parameters from the transformed images $D(\xi)$ and $D_d(\xi)$ matches these arrays to one of a set of arrays of precalculated OTF's by minimizing one of the above norms over the set of arrays. The value of piston used to compute these matched OTF's is then the estimated piston misalignment of the experimental data. Using a previously developed model [8], I computed pairs of diversified OTF's for 100 piston values from -0.49 to 0.50 in steps of 0.01 waves of piston. The smallest norm arising from this set of piston values

yields the estimated piston. As mentioned above, the norms are computed over only the overlap regions of Figure 1, in contrast to full rectangular region used by Ling. Actually, because of the symmetry of the Fourier transform of real data, only one of the two overlap regions is needed and used.

The results of this section are from three data sets (Data1, Data2, and Data3) of detected diversity images of three objects--the first an experimental point source and the other two different views of the same bar chart slide. Since a pair of diversity images could not be experimentally aligned with respect to each other, Ling aligned the data by correlation. The steps used to obtain piston misalignment in the experimental apparatus have values on the horizontal axis of the following three figures. The solid lines graph what the experimental waves piston misalignment should be as a function of these steps. The dashed lines show the values of piston misalignment Ling estimated for the data sets with conjugate-gradient optimization over the full frequency domain. If not otherwise indicated, she used the Wiener-Helstrom norm with $\sigma=0.00005$. All other data in the figures result from discrete optimization. Discrete optimization with the Wiener-Helstrom norm with $\sigma=0.00005$ gives the same results as does optimization with the Gonsalves norm.

Figure 2 shows a comparison of the various optimizations for Data1, the experimental point source. Not shown is the Gonsalves norm optimized over only the horizontal axis of the overlap region, since this is essentially the same as that over the entire overlap region and thus not as good as the vertical axis optimization. The conjugate-gradient optimization is acceptably within the range of error of ± 0.1 wave of piston, while the discrete optimization with the Gonsalves norm has one value with an error of $+0.14$ and another with an error of -0.13 . Using the vertical axis instead of the entire overlap region brings the error within the acceptable tolerance. That the discrete optimization with the Gaussian norm gives the best estimate of piston for a point source is expected, since the ratio of OTF's is the ratio of transforms of point source images. Such a good fit

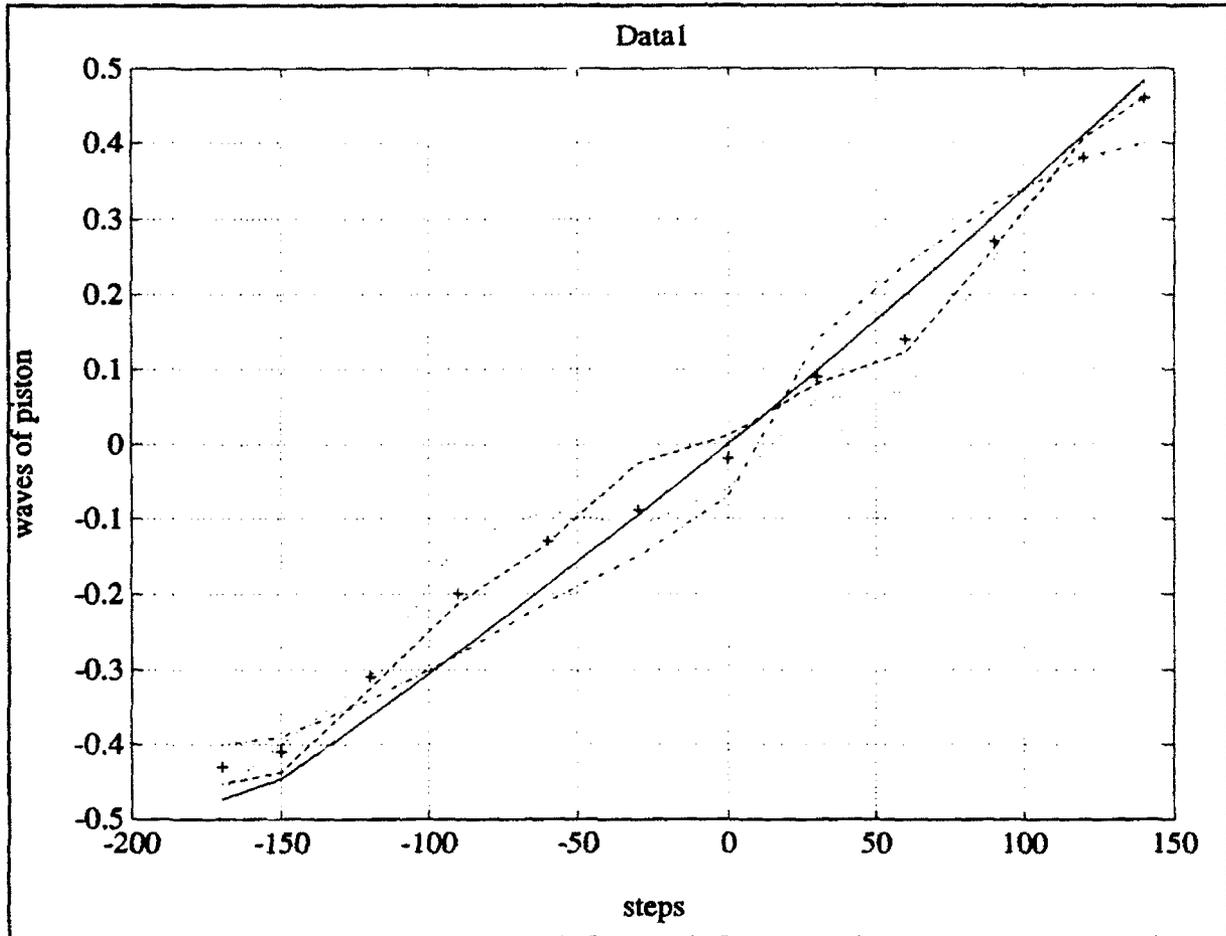


Figure 2. The dot-dashed line results from the Gaussian norm over one overlap region. The dotted line results from the Gonsalves norm over one overlap region, while the "+" data is over the vertical center line of this region.

cannot be expected for Data2 and Data3, which are taken for extended objects.

Figure 3 shows the Euclidean norm to be an extremely poor predictor of piston misalignment. The best estimator of piston is the conjugate-gradient optimization with the discrete Gonsalves norm practically identical to it except at one data value. The discrete estimation at this one value has an error of -0.1375 waves of piston. Again, using the vertical axis of instead of the entire overlap region, brings all errors within the acceptable tolerance of ± 0.01 waves of piston.

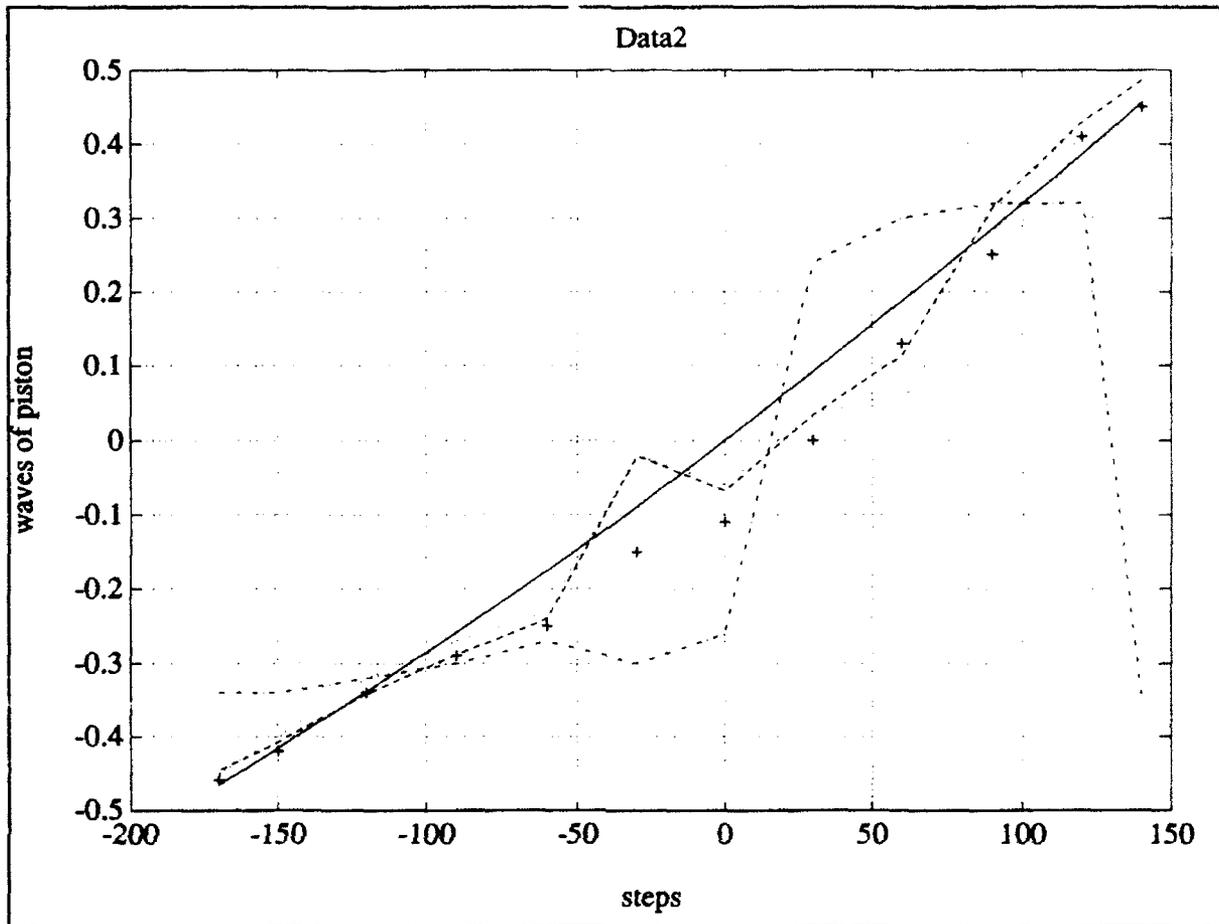


Figure 3. The dot-dashed line results from the Gaussian norm over one overlap region. The dotted line results from the Gonsalves norm over one overlap region, while the "+" data is over the vertical center line of this region.

None of the optimization methods give acceptable estimates of the piston misalignment of Data3, as can be seen in Figure 4. The lower graph of the figure shows that significant improvement may be obtained by using the large smoothing parameter $\sigma=0.5$ in the Wiener-Helstrom norm. Even so, two of the data values have unacceptable piston estimations for the discrete and three values do for the conjugate-gradient optimizations--errors of 0.1625 and 0.1843 for the first and 0.1067, 0.2243, and 0.1262 for the second.

Miller [6] suggests that the problem with the third data set as compared with the second is that it does not have sufficient frequency information. The next section further investigates the differences between these data sets by reconstructing their objects and estimating noise.

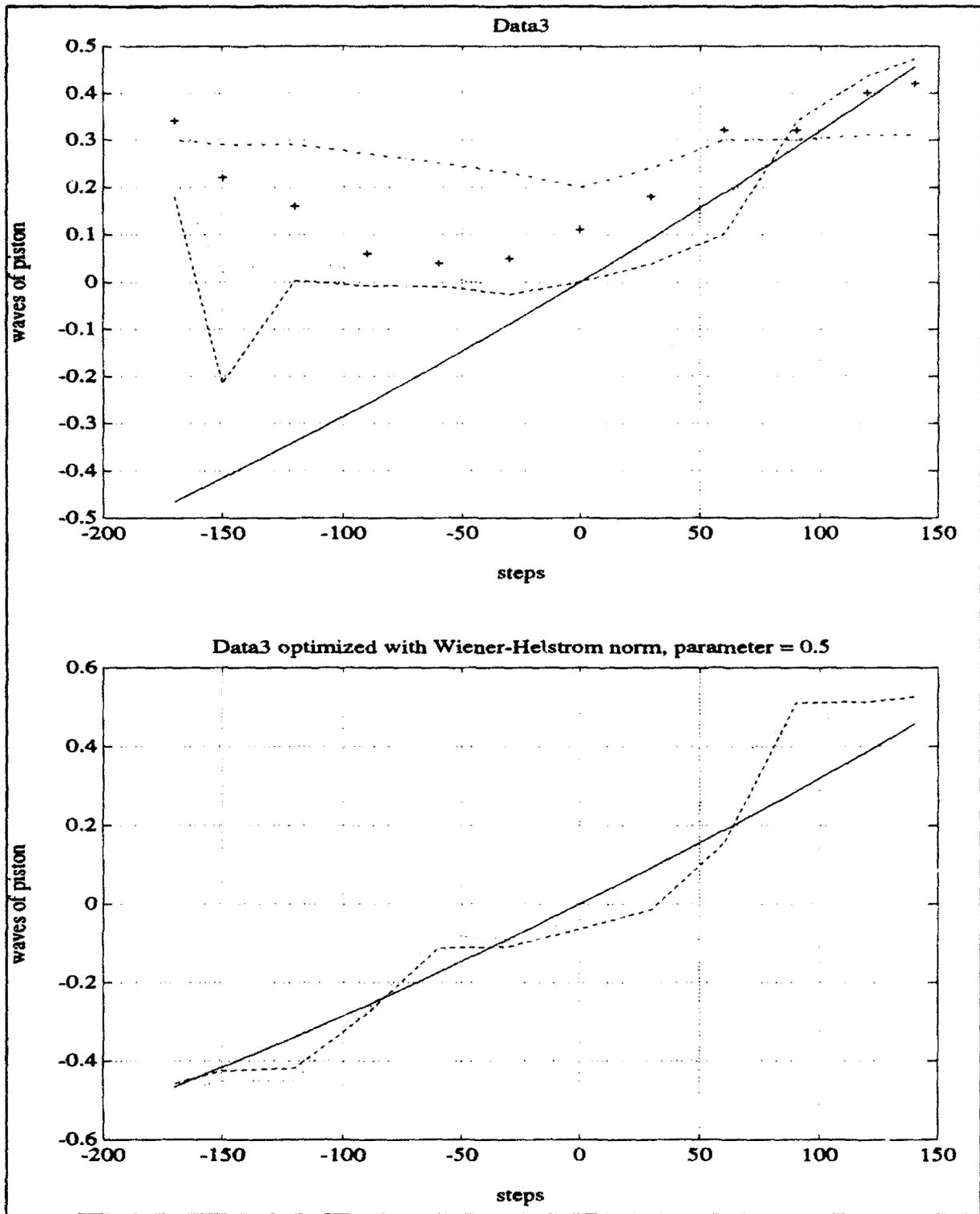


Figure 4. The dot-dashed line results from the Gaussian norm over one overlap region. The dotted line results from the Gonsalves norm over one overlap region, while the "+" data is over the vertical center line of this region.

III. OBJECT RECONSTRUCTION

The Wiener-Helstrom filter, Eqs. (6) and (7) with $\sigma=0.00005$, gives a reconstruction of the Fourier transform of the object for each of the twelve piston data pairs $D(\xi)$ and $D_c(\xi)$ contained in Data2. The inverse transform of the average of these twelve arrays $F(\xi)$ gives the reconstruction of the intensity of the bar chart object of Data2 shown in Figure 5. The intensity values are not as smooth as they appear since the graph shows only every fourth pixel in both horizontal and vertical directions of the 256x256 grid.

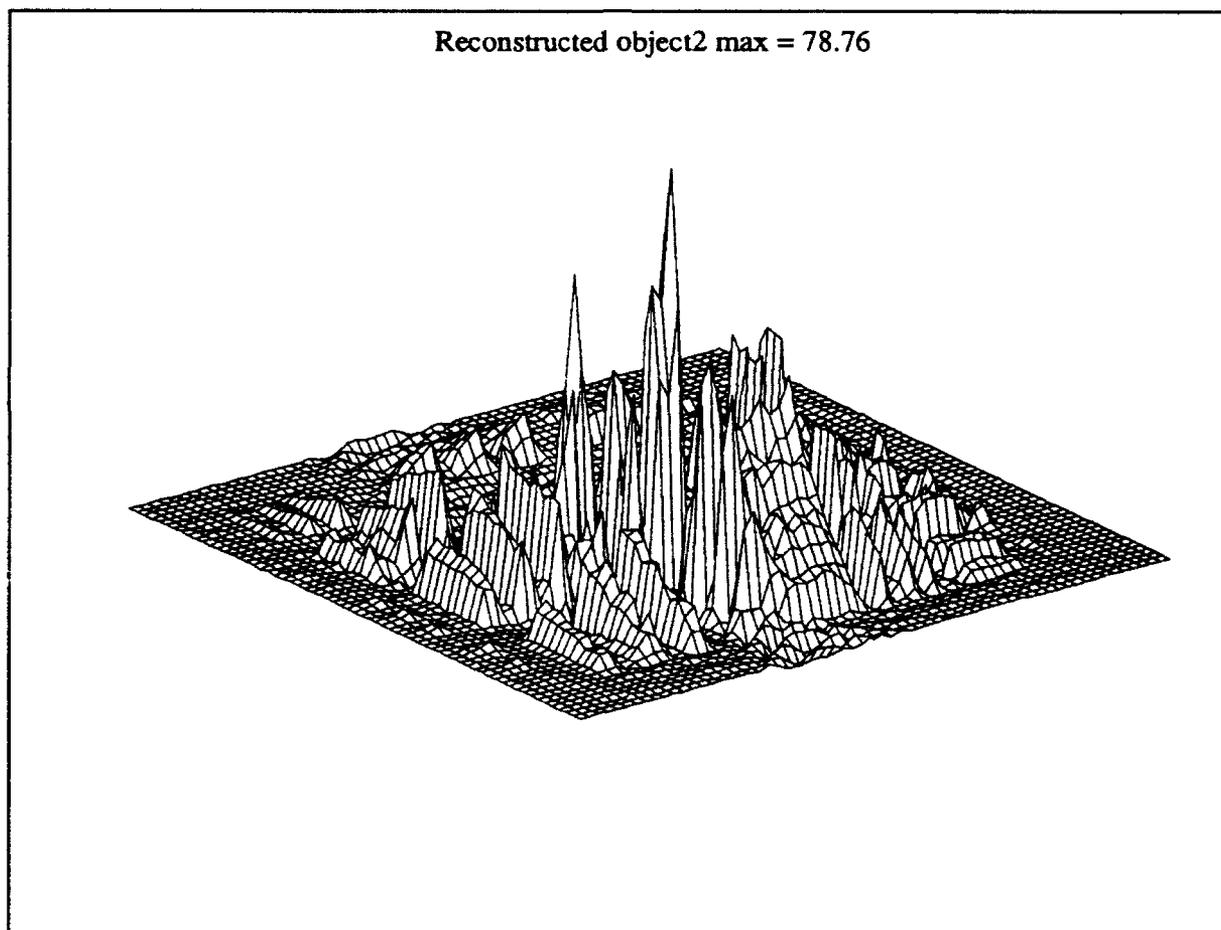
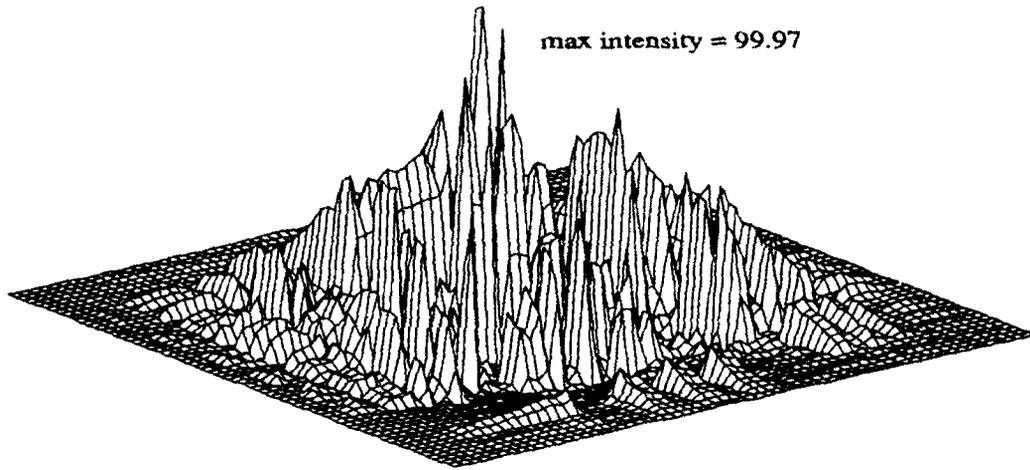


Figure 5. Intensity of the reconstructed object of Data2.

Reconstruction of the object of Data3 is shown for two values of the smoothing parameter σ in Figure 6. The larger value of σ gives a much smoother object but this is at the expense of the noise of Eq. (3). The statistical properties of the noise, obtained from the inverse transform of

Reconstructed object3, parameter = 5.e-5



Reconstructed object3, parameter = 5.e-1

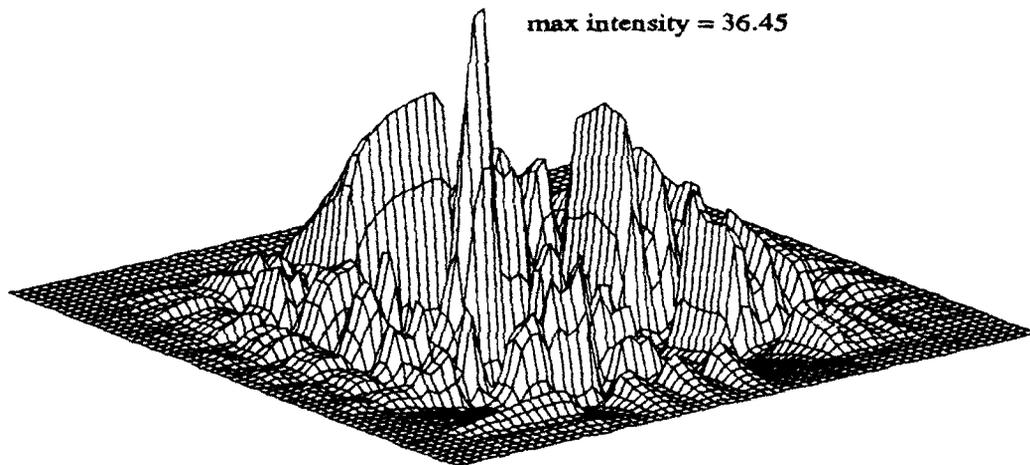


Figure 6. Intensity of the reconstructed object of Data3 for two values of the smoothing parameter σ .

$N(\xi)$ in Eq. (3), average noise \bar{n} and \bar{n}_r , and their standard deviations σ and $(\sigma)_r$, for the diversity data pairs, appear in Table I for Data2 and in Table II for Data3 with $\sigma = 0.00005$ and in Table III for Data3 with $\sigma = 0.5$. Notice that even with the same smoothing parameter, Data3 has more noise than Data2. The object reconstructed for Data3 with the larger value of σ has significantly more noise than with the smaller value. Sample

Table I. Data2

piston	\bar{n}	\bar{n}_r	σ	$(\sigma)_r$
-0.465	0.020	0.004	1.397	1.929
-0.416	-0.007	0.008	1.398	1.857
-0.339	-0.009	0.050	1.499	1.778
-0.259	-0.026	0.059	1.315	1.704
-0.175	-0.017	0.007	1.341	1.727
-0.089	-0.012	0.044	1.254	2.517
0.0	-0.012	0.006	1.189	1.553
0.092	-0.005	-0.031	1.371	1.621
0.188	-0.058	0.023	1.048	2.556
0.286	0.008	-0.080	1.155	2.599
0.387	-0.043	0.035	1.302	1.838
0.456	-0.023	0.066	1.433	1.915

histograms are shown in Figure 7 and in Figure 8.

Using the reconstructed objects and noise would be a way to test telescope models more realistically than with simple objects and Gaussian noise. The next section extends the two circle aperture model of [8] to an arbitrary number of circles. This would represent a more realistic telescope.

Table II. Noise for Data3 with $\sigma = 0.00005$

piston	\bar{n}	\bar{n}_s	σ_s	(σ_s)
-0.465	-0.018	0.009	2.950	4.140
-0.416	-0.037	-0.069	2.673	3.985
-0.339	-0.139	0.101	2.598	4.183
-0.259	-0.076	0.120	2.659	3.996
-0.175	-0.037	-0.040	2.499	3.554
-0.089	0.745	0.142	2.485	3.154
0.0	0.048	-0.030	2.496	2.606
0.092	0.071	-0.054	2.140	2.532
0.188	-0.073	0.150	1.951	2.798
0.286	-0.085	-0.048	2.136	3.576
0.387	-0.237	0.082	2.291	3.398
0.456	-0.064	-0.003	2.513	3.661

Table III. Noise for Data3 with $\sigma = 0.5$

piston	\bar{n}	\bar{n}_s	σ_s	(σ_s)
-0.465	-0.184	-0.157	10.456	9.115
-0.416	-0.203	-0.235	9.972	8.561
-0.339	-0.305	-0.065	9.570	9.311
-0.259	-0.242	-0.046	10.520	9.586
-0.175	-0.203	-0.206	10.542	8.796
-0.089	-0.091	-0.024	11.039	8.825
0.0	-0.118	-0.196	11.444	8.144
0.092	-0.095	-0.219	10.878	7.416
0.188	-0.239	-0.016	9.474	8.767
0.286	-0.251	-0.214	9.154	8.252
0.387	-0.190	-0.084	9.720	8.315
0.456	-0.230	-0.168	8.741	7.802

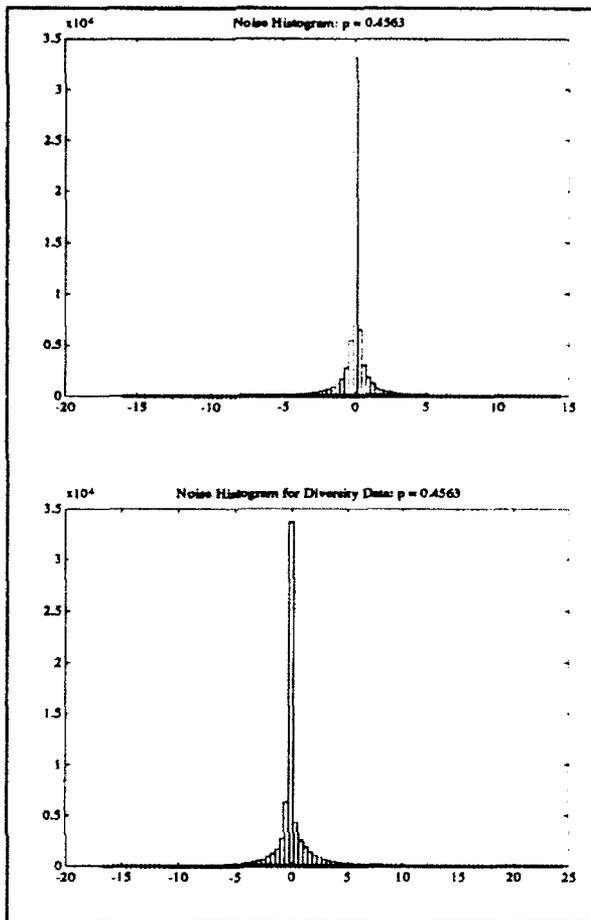


Figure 7. Histogram of noise from the reconstructed object of Data2 for the diversity pair with experimental piston of 0.456, $\sigma=0.00005$.

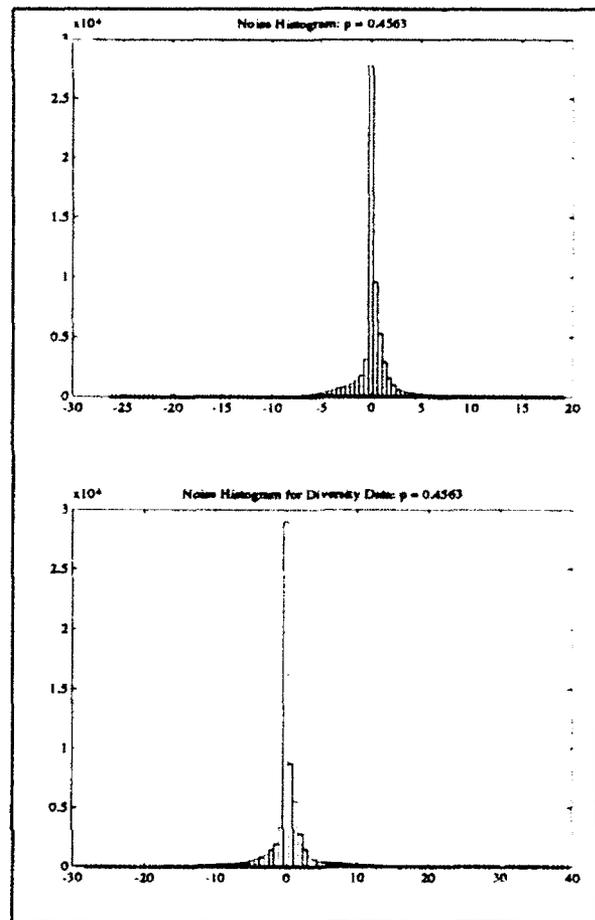


Figure 8. Histogram of noise from the reconstructed object of Data3 for the diversity pair with experimental piston of 0.456, $\sigma=0.00005$.

IV. MANY MIRROR MODEL

The mirror may consist of an even number of circular apertures of common radius r centered at C_k at the same distance from the origin. Thus the pupil function with linear aberrations is

$$P(\mathbf{X}) = \sum_{k=1}^n \text{circ}\left(\frac{\mathbf{X}-\mathbf{C}_k}{r}\right) e^{j(p_k \cdot \mathbf{T}_k \cdot \mathbf{X})} \quad (9)$$

where p_k are piston parameters and \mathbf{T}_k are tilt vectors, and $\text{circ}(\mathbf{X}) = 0$ if $|\mathbf{X}| > 1$ or $= 1$ otherwise. For phase diversity, the image distance is defocused from d , to $d+d$. Equation (5-28) on p.92 of [10] requires the pupil function (9) to be

multiplied by

$$\exp(j\lambda\pi\{1/(d_1+d_2)-1/d_1\}|\mathbf{x}|^2) = \exp(-j\pi d|\mathbf{x}|^2/(\lambda d_1')).$$

Thus, the coherent transfer function is

$$\begin{aligned} H(\mathcal{U}) &= P(-\lambda f l \mathcal{U}) \\ &= \left[\sum_k \text{circ}\left(\frac{-\lambda f l \mathcal{U} - C_k}{r}\right) e^{j(p_k - \lambda f l T_k \mathcal{U})} \right] e^{-j\pi \lambda f l |\mathcal{U}|^2} \\ &= \left[\sum_k \text{circ}\left(\frac{\mathcal{U} + \Gamma_k}{\rho}\right) e^{j(p_k - T_k \mathcal{U})} \right] e^{-j\delta |\mathcal{U}|^2/4}, \end{aligned} \quad (10)$$

where the image distance d_1 satisfies the formula $1/d_1 + 1/d_2 = 1/\ell$ where d_2 is the object distance, which we assume to be infinite, and ℓ is the focal length, the wavelength is λ , $\Gamma = C/(\lambda \ell) = (\gamma, 0)$ and $\rho = r/(\lambda \ell)$, $T_k = \lambda \ell T_k$, $\delta = 4\pi \lambda d$.

The optical transfer function is formed from the products of $H(\Xi + \mathcal{F}) H^*(\Xi - \mathcal{F})$, where * indicates complex conjugate. The exponential factors outside the brackets of Eq. (10) of these products simplify to

$$\exp[-j\delta(|\Xi + \mathcal{F}|^2 - |\Xi - \mathcal{F}|^2)/4] = \exp[-j\delta \mathcal{F} \cdot \Xi]$$

If we let $T_u = T_1 - T_2$ and $p_u = p_1 - p_2$, the unnormalized optical transfer function S at frequency $2\mathcal{F} = (f_x, f_y)$ is

$$\begin{aligned} S(2\mathcal{F}) &= \iint_{\mathcal{X}} H(\mathcal{X} + \mathcal{F}) H^*(\mathcal{X} - \mathcal{F}) d\mathcal{X} = \\ &= \sum_I \sum_K e^{j[p_{IK} - (T_I - T_K) \cdot \mathcal{F}]} \iint_{\mathcal{X}} \text{circ}\left(\frac{\mathcal{X} + \Gamma_I + \mathcal{F}}{\rho}\right) \text{circ}\left(\frac{\mathcal{X} + \Gamma_K - \mathcal{F}}{\rho}\right) e^{j[(T_{IK} - \delta \mathcal{F}) \cdot \mathcal{X}]} d\mathcal{X}, \end{aligned} \quad (11)$$

To evaluate each of the terms in the sum, let $\Xi = \mathcal{X}' - (\Gamma_I + \Gamma_K)/2$:

$$e^{j[p_{IK} - (T_I - T_K) \cdot \mathcal{F} + (T_{IK} - \delta \mathcal{F}) \cdot (\frac{\Gamma_I + \Gamma_K}{2})]} I(\delta, \mathcal{F}, \Gamma_{IK}, \rho, T_{IK}), \quad (12)$$

where $\Gamma_u = \Gamma_I - \Gamma_K$ and

$$I(\delta, F, \Gamma_{ik}, \rho, T_{ik}) = \iint_{\mathbb{R}'} \text{circ}\left(\frac{\mathbb{R}' + \mathbf{A}}{\rho}\right) \text{circ}\left(\frac{\mathbb{R}' - \mathbf{A}}{\rho}\right) e^{-j(\mathbf{T}_{ik} + \delta F) \cdot \mathbb{R}'} d\mathbb{R}', \quad (13)$$

where $\mathbf{A} = \mathbf{F} + \Gamma_{ik}/2$. Integration of the above depends upon whether or not $\mathbf{A} = 0$.

Case 1: $\mathbf{A} = 0$.

$$I(\delta, F, -2F, \rho, T_{ik}) = \int_{|\mathbb{R}'| \leq \rho} \int e^{-j(\mathbf{T}_{ik} + \delta F) \cdot \mathbb{R}'} d\mathbb{R}' \quad (14)$$

which becomes after appropriate variable changes

$$\begin{aligned} I &= \int_{-p}^p e^{-j(\mathbf{T}_{ik} + \delta F)_x \xi} d\xi \int_{-\sqrt{\rho^2 - \xi^2}}^{\sqrt{\rho^2 - \xi^2}} e^{-j(\mathbf{T}_{ik} + \delta F)_y \eta} d\eta \\ &= \frac{4}{(\mathbf{T}_{ik} + \delta F)_y} \int_0^p \cos[(\mathbf{T}_{ik} + \delta F)_x \xi] \sin[(\mathbf{T}_{ik} + \delta F)_y \sqrt{\rho^2 - \xi^2}] d\xi, \end{aligned} \quad (15)$$

provided $(\mathbf{T}_{ik} + \delta F)_y \neq 0$. If this expression equals zero, the integral may be evaluated by taking the limit as it approaches zero. (See [8].)

Case 2: $\mathbf{A} \neq 0$

The transformation

$$\begin{aligned} \xi' &= \mathbf{A}_x \xi - \mathbf{A}_y \eta \\ \eta' &= \mathbf{A}_y \xi - \mathbf{A}_x \eta \end{aligned}$$

moves the circle centers from $\pm \mathbf{A}$ to $(\pm 1, 0)$ and stretches their radii from ρ to $\rho_1 = \rho/|\mathbf{A}|$ and gives a factor to the new area differential of $|\mathbf{A}|^2$.

The circles would not intersect if $\rho_1 \leq 1$, i.e. if $\rho \leq |\mathbf{F} + \Gamma_{ik}/2|$. Thus, we get a contribution from the (i, k) term only if \mathbf{F} lies in the circle of radius ρ about $\Gamma_{ik}/2$. Thus,

$$I = 0 \text{ if } \rho_1 \leq 1.$$

We now assume that $\rho_1 \geq 1$. Thus,

$$I(\delta, F, \Gamma_{12}, \rho, T_{12}) = \frac{\rho^2}{\rho_\lambda^2} \iint_{\mathbf{R}} \text{circ}\left(\frac{\mathbf{R}+(1,0)}{\rho_\lambda}\right) \text{circ}\left(\frac{\mathbf{R}-(1,0)}{\rho_\lambda}\right) e^{-j(\alpha t + \beta \eta)} d\mathbf{R}, \quad (16)$$

where

$$\alpha = -(T_{12} + \delta F) \cdot \mathbf{A}_1 - (T_{12} + \delta F) \cdot \mathbf{A}_2$$

$$\beta = (T_{12} + \delta F) \cdot \mathbf{A}_1 - (T_{12} + \delta F) \cdot \mathbf{A}_2$$

This integral may be simplified to

$$I = \frac{4\rho^2}{\alpha\rho_\lambda^2} \int_0^{\sqrt{\rho_\lambda^2-1}} \cos(\beta\eta) \sin[\alpha(\sqrt{\rho_\lambda^2-\eta^2}-1)] d\eta, \quad (17)$$

which may be evaluated for the special case of $\alpha=0$ by taking the limit.

We have reduced all the integrations to one dimension, which may be evaluated numerically to obtain the OTF S in Eq. (11).

V. PARAMETER REDUCTION

The experiment on which this report is based took care to insure that the two mirrors had no tilt misalignment. Thus, there was only one parameter to estimate, p_{12} , the difference between the pistons of the two mirrors. In practice, each mirror would also have two tilt parameters--the horizontal and vertical components of their tilt, T_x and T_y , respectively. Referring to Figure 1 it can be seen from Eq. (11) the OTF evaluated at $2F$ is

$$\begin{aligned} & e^{-j(p_{12}+(T_{11}+T_{21}) \cdot F)} I(\delta, F, \Gamma_{21}, \rho, T_{21}), \quad \text{if } F \text{ is in circle 1} \\ & e^{-j(-p_{12}+(T_{11}+T_{21}) \cdot F)} I(\delta, F, \Gamma_{12}, \rho, T_{12}), \quad \text{if } F \text{ is in circle 2} \\ & 2e^{-j(T_{11}+T_{21}) \cdot F} \cos[(T_{12}+\delta\Gamma) \cdot F] I(\delta, F, 0, \rho, 0), \quad \text{if } F \text{ is in circle 3.} \end{aligned} \quad (18)$$

A diversified pair of OTF's, S and S_r , is found by taking $\delta=0$ and $\delta \neq 0$, respectively, in Eq. (18). For F along the horizontal axis of Figure 1 near the origin (in circle 3 but not circles 1 or 2), the ratio of the OTF's depends only the horizontal component of the difference of the two tilts. Thus, this line segment could theoretically be used to determine this one misalignment parameter for detected image data. With this value determined,

the other non-overlap regions of Figure 1 could be used to find the vertical component of tilt difference and then improve the estimate of its horizontal component. With these two parameters determined, the overlap region could then be used to determine the piston difference, p_{11} . From the ratio of diversified OTF's, only the relative piston and tilt can be estimated but this is sufficient.

The situation for more than two mirrors is much more difficult. It is impossible to reduce the problem to regions in which one parameter may be found in sequence. Theoretically, however, the problem can be reduced to finding the inverse function of one (difference of two pistons) or two (the horizontal and vertical components of the difference or sum of two tilts) parameters by selecting subregions in the proper order. As in the two mirror case, the region around the origin is critical. That this region becomes small and leaves little room for noise can be seen from the geometries of four (Figure 9) and six mirror (Figure 10) problems.

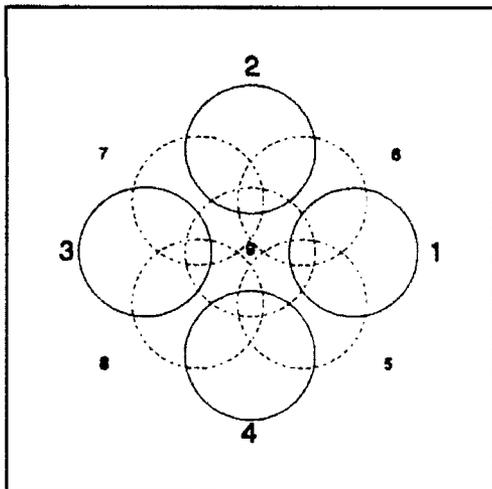


Figure 9. The dark circles represent four mirror apertures. The other circles are also needed for different computations of the OTF.

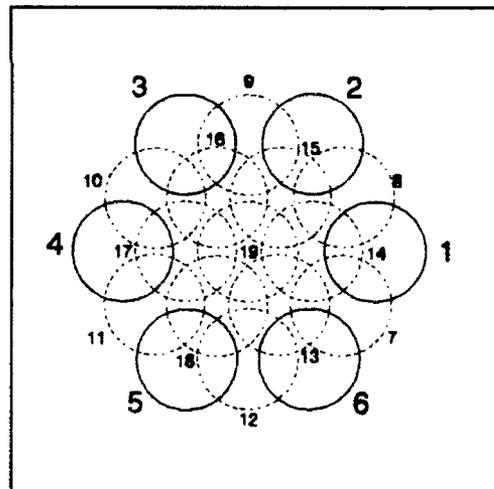


Figure 10. The dark circles represent six mirror apertures. The other circles are also needed for different computations of the OTF.

VI. CONCLUSIONS

The model developed in [8] has now been validated and shown to estimate piston misalignment in an experimental two mirror telescope.

Extension to more than two mirrors with piston and tilt misalignments is much more difficult because of the need to find the inverse values of a two-variable function rather than one and because of the small size of the regions which allow for the reduction of the many variable optimization problem down to one or two variables. The method developed in this report should be tried out on a four or six mirror telescope. Using the object and noise found in Sect. III would provide a more realistic test than a simple object with Gaussian noise.

VII. REFERENCES

1. Gonsalves, R. A., "Phase retrieval and diversity in adaptive optics", Optical Engineering, 1982, Vol. 21, No. 5, pp. 829-832.
2. Paxman, R. G. and Fienup, J. R., "Optical misalignment sensing and image reconstruction using phase diversity", J. Opt. Soc. of Am. A, 1988, Vol. 5, No. 6, pp. 914-923.
3. Paxman, R. G. and Crippen, S. L., "Aberration correction for phased-array telescopes using phase diversity", S.P.I.E., 1990, Vol. 1351, Digital Image Synthesis and Inverse Optics.
4. Paxman, R. G., et al., "Joint estimation of object and aberrations by using phase diversity", J. Opt. Soc. of Am. A, 1992, Vol. 9, No. 7, pp. 1072-1085.
5. Angel, R. P., et al., "Adaptive optics for array telescopes using neural-network techniques", Nature, 1990, Vol. 348, pp. 221-224.
6. Miller, Nancy, LIMI, Phillips Laboratory, Kirtland Air Force Base, NM 87117.
7. Ling, Alice, BDM International Inc., 1801 Randolph Road S.E., Albuquerque, NM 87106.
8. Schruben, J. S., "Misalignment parameters of a double aperture telescope obtained as a function of the ratio of optical transfer functions without and with diversity", Final Report, 1990-91 Research Initiation Program, December 30, 1991, A.F.O.S.R. contract no. F49620-88-C-0053/SB5881-0378.
9. Gonsalves, R.A. and Childlaw, R., "Wavefront sensing by phase retrieval", in Applications of Digital Image Processing III, A.G. Teacher, ed., Proc. Soc. Photo-Opt. Instrum. Eng., 1979, Vol. 207, pp. 32-39.
10. Goodman, J.W., Introduction to Fourier Optics, New York, McGraw-Hill Book Co., 1968.

SATELLITE FRAGMENTATION DUE TO EXPLOSION AND COLLISION

A. Tan
Professor
Department of Physics

Alabama A & M University
Meridian Street
Normal, AL 35762

Final Report For:
Summer Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

August 1992

SATELLITE FRAGMENTATION DUE TO EXPLOSION AND COLLISION

A. Tan
Professor
Department of Physics
Alabama A & M University

Abstract

Past studies to determine the cause of satellite fragmentations in space have been limited by the phenomenological nature of the investigations. In this study, a more analytical approach has been initiated by noting the physical processes taking place in an explosion and a hypervelocity collision. It is argued that an explosion is more effective in producing trackable fragments than a collision with the same amount of available energy. This assertion is substantiated by investigating an on-orbit satellite explosion as exemplified by the NOAA-3 Rocket explosion in 1973 and a hypothetical event in which a similar rocket is impacted by a projectile at 7 km/s. An empirical code (IMPACT 2.0) and an analytical code (2-d MAGI) have been employed to attain the objective. While IMPACT 2.0 modelled the explosive event reasonably well, it failed to fragment the satellite through hypervelocity impact. The mass (and therefore energy) of the projectile had to be increased tenfold before fragmentation could be achieved. The MAGI code modelled the ejecta from the collision event very precisely as expected, but fell short of fragmenting the satellite. At present, both the empirical and analytical codes need significant upgrading before satellite fragmentation events could be analyzed accurately.

SATELLITE FRAGMENTATION DUE TO EXPLOSION AND COLLISION

A. Tan

INTRODUCTION

Since the breakup of Transit 4A rocket in 1961, over one hundred artificial earth satellites have fragmented in orbit, creating a large amount of hazardous material in space. Most were blown up by explosion (whether inadvertent or deliberate) and some are known or alleged to have been broken up by hypervelocity impact as part of anti-satellite (ASAT) testing. Yet others are classified as belonging to the unknown category. To determine the cause of a fragmentation event in space from ground-based observations remains a lingering problem for space debris researchers and strategic planners alike. To date, significant progress in this area have been registered. From ground-based experiments, Bess (1975) observed the characteristic differences between the debris created by explosion and by hypervelocity impact. Culp and McKnight (1986) applied this difference to study several cataloged satellite fragmentation events and attempted to determine the probable cause of fragmentation of events belonging to the unknown category. They used the cumulative mass and velocity perturbation distributions based on Bess's (1975) study as the principal criteria to distinguish between high-intensity explosion, low-intensity explosion and collision-induced breakups. Badhwar, et al. (1988) added the radar cross-section and plane change angle distributions to achieve the same. Tan (1987) showed that the traditional equations used to calculate the velocity perturbations of the fragments (Culp and McKnight, 1986; Kling, 1986) broke down in five different cases, some of which were closely approximated in several actual fragmentation events. Badhwar, et al. (1990) subsequently developed alternative equations for velocity perturbations, which are applicable for all occasions.

Inspite of the substantive progress made in the orbital debris study, determining the probable cause of fragmentation remains handicapped at the present time. The major difficulty is the inability of the tracking system to observe small particles (smaller than about 10 cm in diameter). For it is in the

mass distributions of these small particles that the major difference between the explosional and collisional signatures lie. The cumulative mass distributions for trackable fragments seem to follow the exponential pattern for both explosion and collision (Johnson and McKnight, 1987). Until the resolution of ground-based radar observations can be improved significantly, the mass distributions from Bess's (1975) observations cannot be applied reliably to distinguish between explosions and collisions. Secondly, the past studies are mostly phenomenological in nature and do not consider the physical processes involved in the fragmentation. In view of the inherent limitations of the phenomenological aspects of the past studies, a new approach to the satellite fragmentation phenomena is initiated in this study. Instead of merely observing the debris characteristics following a fragmentation, we shall, in addition, take note of the physical processes taking place during the fragmentation.

OBJECTIVES

It must be stated at the outset that both explosions and collisions contain different scenarios which must be treated separately. For example, an explosion can be physical (e.g. a pressure burst), chemical or even nuclear. Further, a chemical explosion may belong to the deflagration (a low-intensity explosion) or detonation (high-intensity explosion) category. Collisions can take place between two objects of comparable size (such as in a planned ASAT test) or between a large object and a small projectile (as in an accidental collision between a spacecraft and an orbital debris). To overlook these considerable differences, we shall take two simplified breakup scenarios: one, a low intensity explosion of a rocket body as exemplified by the Delta second stage rocket body explosions, and two, a hypervelocity impact of a comparable satellite by a projectile. To compare the breakup efficiency in each case, we assume that both processes release the same amount of energy. This assumption will ensure that the projectile belongs to a small object category, e.g. an orbital debris. Based on the first principles and the physical nature of explosions and hypervelocity impact, we can qualitatively predict the general characteristics of the fragments

produced in the two cases. We shall validate these predictions by running existing numerical codes available at the Phillips Laboratory. Our final objective is to determine if the results can be applied to delineate actual satellite fragmentations in space produced by the two distinct causes.

THEORY

In a chemical explosion, the energy is released by two exothermic reactions in succession, the first being the reduction of nitrate groups to form nitrogen gas and the second, the oxidation of carbon and hydrogen to form carbon dioxide and water vapor (cf. Marshall, 1989). The rapidly expanding gas creates a blast wave which exerts sudden intense pressure on the walls of the container and if the stress exceeds the mechanical strength of the material, cracks form and propagate through the material, consequently fragmenting the container. In a low intensity explosion involving propellants, the maximum pressure is usually limited to 340 bars and temperature to 2400C (Holmes, 1992). Approximately 50% of the energy released is carried away by the escaping gases and heat conduction. Assuming that the velocity of the explosive gases varied linearly with the radius and applying the energy conservation principle, Gurney (1943) derived the velocity (assumed constant) of the fragments of exploding spherical and cylindrical shells. Taylor (1963) deduced the angle of a fragment from the normal to the surface of an exploding cylinder. Anderson, et al. (1985) showed that better results were obtained if one incorporated gas leakage in the model.

In a hypervelocity impact involving a target and a projectile, a large part of the kinetic energy is turned into intense localized heat which produces partial vaporization, spallation and fracture. At a relative velocity of about 10 km/s, the instantaneous pressure at the point of contact is estimated to reach 5 megabars (Bjork and Olshaker, 1965). The remaining energy produces transverse stress far exceeding the material strength and tolerance of the material and it is this late time effect that causes the breakup producing secondary spray particles and fragments. In laboratory experiments with hypervelocity impact in the 3-4.5 km/s range, over 30% of the mass were not recovered and presumably

turned into vapor (Bess, 1975).

The differences between explosion and hypervelocity impact are indeed fundamental in nature. First, the initial energy in an explosion is chemical potential energy whereas in a hypervelocity impact, it is purely translational kinetic energy (rotational kinetic energy, if any, can be shown to be negligible compared with translational). Second, this energy is applied externally in a hypervelocity impact, whereas it is applied internally in an explosion [an explosion external to a satellite is possible in a planned experiment in space (Aviation Week and Space Technology, 1986), but such a special case is not considered here]. Third, the energy is initially applied locally on a relatively small area in an impact (for a small projectile) whereas it is distributed over more uniformly over a larger area in an explosion (unless the charge is in direct contact with the wall of the container). Based on purely physical considerations, one would expect a small projectile collision to be less effective in creating fragments than an explosion with the same amount of energy. For, in a collision, a large amount of projectile energy is expended in melting, vaporization, spallation, target penetration and perforation and energy dissipation in the form of heat and intense optical radiation. If the energy of the projectile is too small, the target may not fragment. Also, at sufficiently high energy, the projectile might actually perforate the wall twice and exit the target, leaving little energy to fragment the structure. In general, a small projectile is not expected to be efficient in creating debris of trackable size.

Normally, an impact creates a large number of small fragment at high speeds and a small number of large fragments with low speeds. The angular distributions of all fragments, particularly the small ejecta will be highly anisotropic. In comparison, an explosion will create a much smaller number of small fragments, a larger number of large fragments with significantly higher velocities and a more isotropic distribution of the fragments [unless the fragmentation is in the "clam" or "half-segment" mode (Benz, et al., 1987), in which case the debris distribution will be anisotropic]. A direct consequence of small velocity perturbations of the larger collisional fragments is small angles of plane change

(Badhwar, et al., 1988), In consequence, we expect much smaller plane change angles for the larger trackable fragments of a collisional breakup (due to a small projectile) in space.

Setting aside the question of angular distribution of the fragments and the deviation of the fragment velocities, we can estimate the spread velocity of the fragments in an isotropic fragmentation from the energy and momentum conservation principles. We shall first consider a satellite fragmentation due to explosion with definite explosive energy. Let M be the mass of the satellite, E_0 the energy liberated in the explosion and W_c the total work done in structural breakup, expulsion of gases, energy dissipation, etc. Then $E_0 - W_c$ is the energy available for spreading the fragments. Let v_r be the radial spread velocity of the fragments (assumed to be the same for all fragments) in a frame of reference moving with the satellite. By the law of conservation of energy,

$$\frac{1}{2} M v_r^2 = E_0 - W_c \quad (1)$$

from which, we get, by retaining the positive root only:

$$v_r = \left[\frac{2}{M} (E_0 - W_c) \right]^{\frac{1}{2}} \quad (2)$$

Next we shall consider the fragmentation of the same satellite in a head-on collision with a small projectile whose translational kinetic energy is equivalent to the explosive energy considered above. If m is the mass of the projectile ($m \ll M$) and v_0 its velocity relative to the target, then

$$\frac{1}{2} m v_0^2 = E_0 \quad (3)$$

We consider the general case in which the projectile perforates both walls and exits the target carrying with it a small mass μ of target material [in general $\mu < m$ (cf. Giere, 1959)]. In this case W_c represents the total work done in melting, vaporization, spallation, perforation and energy dissipation besides structural fragmentation, and is clearly greater than W_c . Assume that the target is broken into n fragments. Let v_i be the exit velocity of the projectile (along

the z-direction), v_r , the radial component of velocity received by a fragment and v_z , the z-component of velocity received by the fragment. Then the resultant velocity of the fragment is given by

$$\vec{v}_f = \vec{v}_r + \vec{v}_z \quad (4)$$

By the law of conservation of linear momentum,

$$m v_o = (M-\mu) v_z + (m+\mu) v_f \quad (5)$$

Thus

$$v_z = \frac{m}{M-\mu} v_o - \frac{m+\mu}{M-\mu} v_f \quad (6)$$

By the law of conservation of energy,

$$\frac{1}{2} m v_o^2 = \frac{1}{2} (m+\mu) v_f^2 + W_c + \sum_{i=1}^n \frac{1}{2} \frac{M-\mu}{n} \vec{v}_i \cdot \vec{v}_i \quad (7)$$

By Eq.(4), the last term on the right hand side expands to

$$\sum_{i=1}^n \frac{1}{2} \frac{M-\mu}{n} \vec{v}_i \cdot \vec{v}_i = \frac{1}{2} (M-\mu) v_r^2 + \sum_{i=1}^n \frac{M-\mu}{n} \vec{v}_r \cdot \vec{v}_z + \frac{1}{2} (M-\mu) v_z^2 \quad (8)$$

The middle term on the right hand side of (8) is equivalent to the flux of a constant vector through a closed surface and sums out to zero. Substituting from Eq.(6), simplifying and applying the approximation $M \gg m \gg \mu$, we get, from the energy equation:

$$\frac{1}{2} m v_o^2 = \frac{1}{2} (M-\mu) v_r^2 + \frac{1}{2} (m+\mu) v_f^2 + W_c \quad (9)$$

whence

$$v_r = \left[\frac{m}{M-\mu} v_o^2 - \frac{m+\mu}{M-\mu} v_f^2 - 2 \frac{W_c}{M-\mu} \right]^{\frac{1}{2}} \quad (10)$$

The spread velocity of the collisional fragments is thus given by equations (6) and (10).

Equation (9) clearly implies that much of the projectile energy is expended in perforation, spallation, etc. Consequently, collision is not an effective mechanism (compared with explosion) in satellite fragmentation and fragment

dispersion. This result has been intuitively assumed in past studies (Culp and McKnight, 1986; Badhwar, et al., 1988; Tan, 1987; Badhwar, et al., 1990) but never shown to be true. If the projectile is completely absorbed by the target, $\mu=0$, $v_r=0$, and

$$v_r = \left[\frac{m}{M} v_o^2 - 2 \frac{W_c}{M} \right]^{\frac{1}{2}} \quad (11)$$

By virtue of Eq.(3), this is of the same form as Eq.(2). However, even in this limiting case (since W_c is usually greater than W_e), collision is not expected to be as efficient in fragment spreading as explosion from the same amount of released energy.

COMPUTATIONS

As stated earlier, two distinct satellite fragmentation scenarios were simulated in this study, one, due to explosion and the other, due to collision. The explosive event was the NOAA-3 Delta 2nd stage Rocket explosion of 1973, which was one of a series of Delta rocket malfunctions in orbit caused by the ignition of excess fuel left in the tank and triggered, most probably, by solar heating (Gumpel, 1982). The collision scenario is a hypothetical event in which a similar rocket is fragmented by hypervelocity impact in orbit. In order to faithfully compare the fragment characteristics in the two cases, the energy released in the two events are taken to be equal. This energy is estimated as follows. The average mass of the Delta 2nd stage rockets which exploded in space is about 900 kg (Badhwar and Anz-Meador, 1989). Also, the weighted average of the spread velocity of the fragments of all Delta class explosions is about 82 m/s (Badhwar, et al., 1990). This means that an amount of energy of approximately equal to 3 MJ was responsible for spreading the fragments in a typical Delta class explosion. Since this can be taken to represent about 50% of the total energy released in the explosion, the latter is taken to be 6 MJ. For the collision scenario, a relative velocity of 7 km/s between the target and the projectile was assumed. The translational kinetic energy of 6 MJ then corresponds to a mass of only 0.25 kg. This ensures that the collision belongs to the head-on

category, i.e., between a large target and a small projectile (Chobotov and Spencer, 1991). Assuming the projectile to be an aluminum sphere, the mass of 0.25 kg translates to a diameter of 5.6 cm.

Two numerical codes belonging to two different categories (viz., empirical and analytical) were employed in this study. The empirical code IMPACT 2.0 was written specifically to model satellite fragmentation events (Chobotov and Spencer, 1991). It was originally written to predict the fragment products of the Solwind P-78 and Delta 180 ASAT tests and later modified after the actual test data. The MAGI Code is one of several analytical codes currently operational at Phillips Laboratory in Kirtland Air Force Base. It is an SPH (Smoothed Particle Hydrodynamics) code based on the Lagrangian (i.e., finite element) scheme. Like other analytical codes, MAGI accurately models the hypervelocity impact phenomenon. Because of time limitation, only the 2-dimensional version of the code was used.

IMPACT 2.0 asks for the 6 orbital elements of the target satellite at the point of fragmentation. These orbital elements were calculated from the satellite fragmentation catalog (Johnson and Nauer, 1987). The detailed results of the output are given by Tan (1992). Here we shall only sketch the highlights. According to IMPACT 2.0, the NOAA-3 Rocket suffered a thorough explosive fragmentation, producing 246 trackable objects (fragments larger than about 10 cm in diameter). This compares with an average of 176.5 fragments (from Johnson and Nauer, 1987) for the NOAA and Landsat series rocket explosions. Considering the fact that a significant number of fragments deorbit rapidly and the data are often several days old, this comparison should be regarded as a favorable one (It may be stated that a retrograde velocity perturbation of only 152 m/s is sufficient to deorbit a fragment from low-earth orbit immediately). Figure 1 shows the mass distribution and the cumulative mass distribution according to IMPACT, the latter being a direct consequence of the former. Both distributions resemble the normal expected patterns. Figure 2 shows the number distribution of the fragments with spread velocity. Also shown in the figure as a smooth curve is the observed beta distribution for the Delta rocket explosions, which is taken

from Badhwar, et al. (1990) and normalized for the number of fragments. If one overlooks the dispersion of the computed data, the agreement with observation appears to be a close one. Since IMPACT assumes a random distribution of fragments in space with pairs in antipodal directions, no angular dependence in the computed fragment distribution is expected. Figure 3 shows the spread velocity as a function of the fragment diameter and consists of two straight lines with negative slopes, the line representing the smaller fragments being steeper than that for the larger fragments. The least-square fit of the observed data for the trackable fragments as shown by the broken line in the figure (from Badhwar and Tan, unpublished) is actually flatter and slightly curved.

In order to compare the two basic fragmenting mechanisms of explosion and hypervelocity collision, the NOAA-3 rocket is next collided with a projectile of mass 0.25 kg at a relative velocity of 7 km/s in a hypothetical event. This ensures that the translational kinetic energy of the projectile is 6 MJ, i.e., equivalent to the energy of explosion considered earlier. For the sake of comparison, we assume that 50% of this energy is transferred to spreading the fragments. However, fragmentation was not achieved by this projectile according to IMPACT. The mass of the projectile (and therefore its kinetic energy) had to be increased tenfold to 2.5 kg before fragmentation was achieved. A simple calculation using the mass and the physical dimensions of a typical 100/2000 series Delta 2nd stage rocket gives an approximate thickness of the rocket (Tan, 1992). If we assume a uniform cylindrical drum of length 6 m, diameter 1.5 m, mass 900 kg and made of aluminum alloy (density 2.7 gm/cc), the thickness turns out to be just under 1 cm. According to earth-based experiments, a complete perforation of this structure should occur for striking velocities greater than 1 km/s (cf. Zukas, et al., 1982, p.173). However, no projectile cloud was created by IMPACT 2.0, suggesting that perforation had not occurred. The reason for this failure is not clear at this time but is being communicated to the authors.

In view of the one order of magnitude higher energy involved in the collision, no faithful comparison of the two fragmentation processes using IMPACT could be made. Nevertheless, we shall briefly discuss the basic differences in

the fragment characteristics as provided by IMPACT 2.0. First, the number of trackable fragments produced by collision was only 26 (compared with 246 for explosion). The total number of all fragments for collision was also much smaller compared with explosion (282 vs. 1170). The number distribution of fragments with mass was triangular as usual, but shifted towards the smaller fragments (Fig.4). The number distribution vs. spread velocity plot did not betray any recognizable pattern at all. There was a concentration of small particles of masses less than 0.01 kg centered around 9 km/s. In general, the spread velocities were higher than those in explosion, but this is to be expected considering the 10 times greater energy involved in the impact. Overall, the computations agree with the general notion that the small projectile collision is far less effective in creating and spreading fragments of trackable size than explosion.

Finally, the impact event was also simulated by running the 2-d MAGI Code. For the sake of simplicity, the NOAA-3 Rocket is approximated by a cylindrical aluminum drum of length 6 m, outer diameter 1.5 m and thickness 1 cm. For the sake of symmetry, we assume that it is impacted normally at its mid-section by an aluminum sphere of diameter 5.6 cm at a relative velocity of 7 km/s. The 2-d MAGI Code was run in two mutually perpendicular planes: one, containing the axis of the drum and the projectile (side view); and two, a plane perpendicular to the axis of the drum and containing the projectile (axial view). Figure 5 shows the side view of the material plot of the impact at intervals of 50 microseconds. In under 200 microseconds, the projectile had perforated through both walls of the satellite, but had become thoroughly splintered in the process. The second hole is many times larger than the first in consequence of this splintering. Figure 6 shows the axial view of the material plot of the impact at the same intervals. The results are quite consistent with those of Fig.5. Only the final damage to the structure appears to be even more extensive. The ejecta formation by the MAGI Code for impact is believed to be accurate as is the case with most successful analytical codes. However, the main structure of the satellite had remained intact through this impact. Further, fragment definition in the model is difficult if not impossible at this time.

SUMMARY AND RECOMMENDATIONS

Satellite fragmentation study is still a phenomenological science at the present time. In this study, two diametrically opposite types of codes (one empirical and the other analytical) were employed to simulate satellite explosions and collisions. Although the results obtained by the two codes were quite different, they are also often complementary. For instance, the empirical code (IMPACT) can predict the larger (trackable) fragments of a fragmentation quite well, but not the finer ejecta. The analytical code (MAGI), on the other hand, models the ejecta quite accurately, but not the larger trackable fragments. The IMPACT Code was specifically written to analyze satellite fragmentations and is better equipped to do so at the present time. There is room for improvement in this code, particularly for small projectile collisions. However, this code will always remain empirical at its foundation. In principle, an analytical code (e.g., MAGI) has greater promise in the long run than its empirical counterpart. But major breakthroughs in two areas will be required to realize this promise. First, a fragment definition scheme needs to be incorporated from the smoothed particle products. Secondly, a structural failure model based on theoretical foundations has to be attained (Zukas, et al., 1982, p.408). Until such breakthroughs are attained, it is recommended that both the empirical and analytical models be refined (within their specific limitations) and employed at the same time as the necessary tools to study, analyze and understand the complex phenomenon of satellite fragmentation.

ACKNOWLEDGEMENTS

The author is grateful to the Air Force Office of Scientific Research and Research and Development Laboratories for selecting him as a Summer Associate. Dr. Firooz A. Allahdadi, Chief of Space Kinetic Impact and Debris Branch suggested the topic of this study as well as provided useful direction for which the author is thankful. Lt. Scott Maethner and Mr. James Winter performed the computations of this study. Their efforts are hereby thankfully acknowledged.

REFERENCES

- C. E. Anderson, W. W. Predebon and R. R. Karpp, *Int. J. Engg. Sci.*, 23, 1317, 1985.
- Aviation Week and Space Technology*, Sept. 15, 1986.
- G. D. Badhwar and P. D. Anz-Meador, *Earth, Moon and Planets*, 45, 29, 1989.
- G. D. Badhwar, A. E. Potter, P. D. Anz-Meador and R. C. Reynolds, *J. Spacecr. Rockets*, 25, 420, 1988.
- G. D. Badhwar, A. Tan and R. C. Reynolds, *J. Spacecr. Rockets*, 27, 299, 1990.
- F. J. Benz, C. V. Bishop and M. B. Eck, *Proc. Upper Stage Breakup Conf.*, Houston, TX, May 1987.
- T. D. Bess, *NASA Langley Res. Center Rept. L-10477*, Dec. 1975.
- R. L. Bjork and A. E. Olshaker, *RAND Corp. Memo. RM-2926-PR*, May 1965.
- V. A. Chobotov and D. B. Spencer, *J. Spacecr. Rockets*, 28, 670, 1991.
- R. D. Culp and D. S. McKnight, *Teledyne Brown Engg. Rept. CR-SC-7460*, Oct. 1986.
- A. C. Giere, *NAVORD Rept. 6621*, 1959.
- C. S. Gumpel, *McDonnell Douglas Rept. MDC-H0047*, 1982.
- R. W. Gurney, *Ballistic Res. Lab. Rept. 405*, 1943.
- H. H. Holmes, *The Encyclopedia Americana*, Vol. 10, Grolier Inc., Danbury, CT, 1992.
- N. L. Johnson and D. J. Nauer, *Teledyne Brown Engg. Rept. CS88-LKD-001*, 1987.
- R. Kling, *Teledyne Brown Engg. Rept. CS86-LKD-001*, 1986.
- A. Marshall, *Dictionary of Explosives*, Gordon Press, 1989.
- N. L. Johnson and D. S. McKnight, *Artificial Space Debris*, Orbit Book Co., Malabar, FL, 1987.
- A. Tan, in *NASA Rept. CR-172009*, 1987.
- A. Tan, *USAF Phillips Lab. Rept. WSSD-92-003*, 1992.
- Selected Papers of G. I. Taylor*, Vol. III, Cambridge, 1963.
- J. A. Zukas, T. Nicholas, H. F. Swift, L. B. Greszczuk and D. R. Curran, *Impact Dynamics*, John Wiley and Sons, New York, 1982.

Fig.1. Mass distribution and cumulative mass distribution of fragments of NOAA-3
Rocket breakup according to IMPACT 2.0.

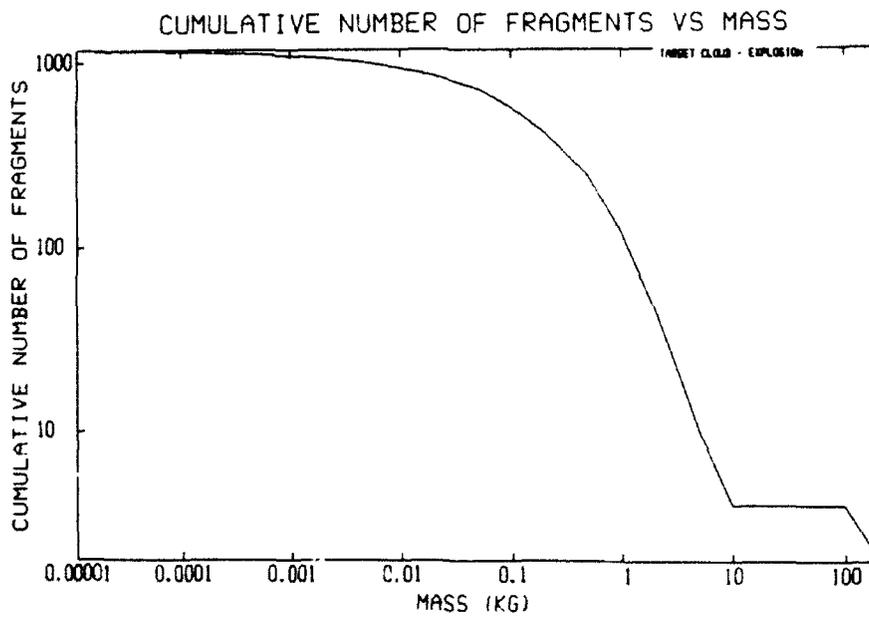
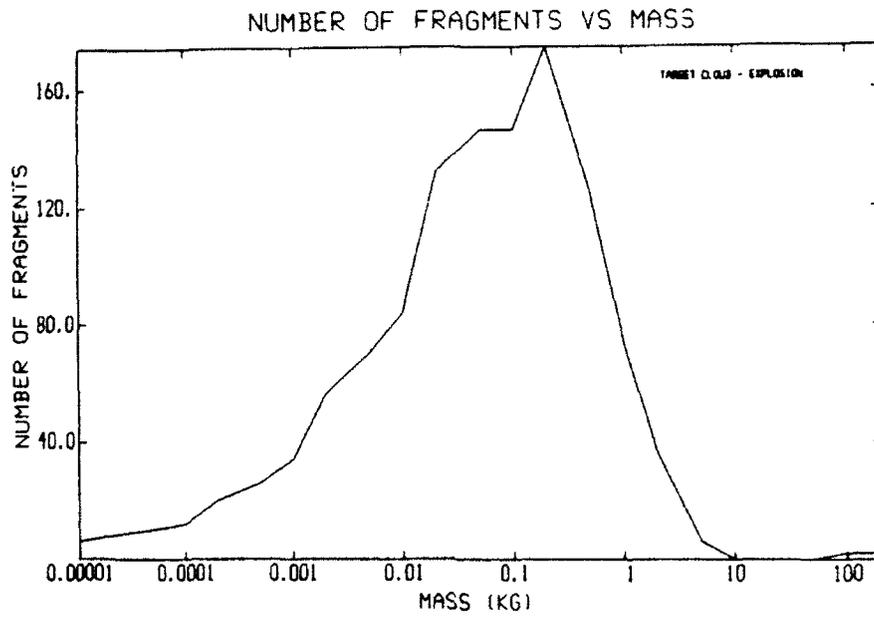


Fig.2. Spread velocity distribution of fragments of NOAA-3 Rocket according to IMPACT 2.0.

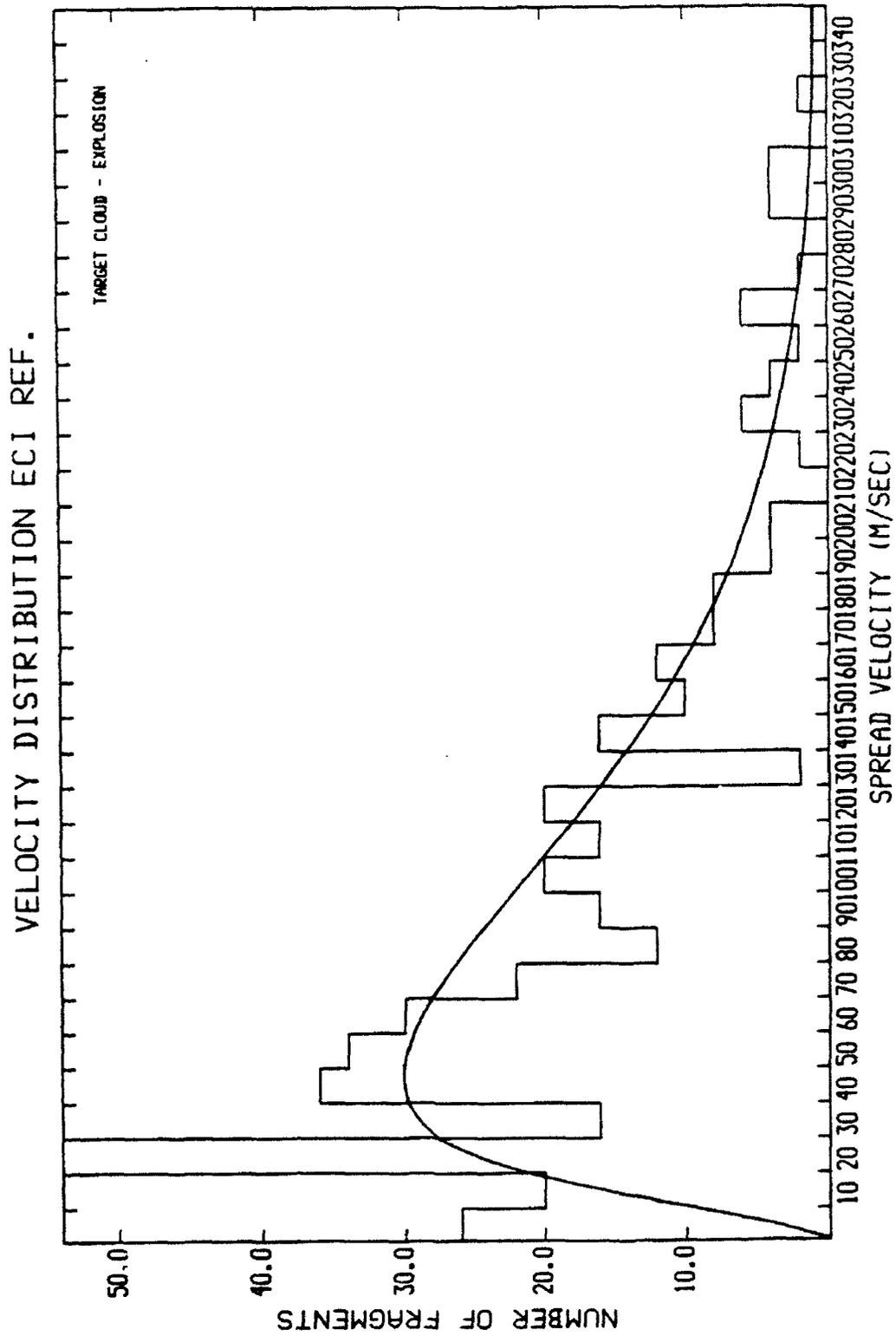


Fig.3. Spread velocity as a function of diameter of fragments of NOAA-3 Rocket according to IMPACT 2.0.

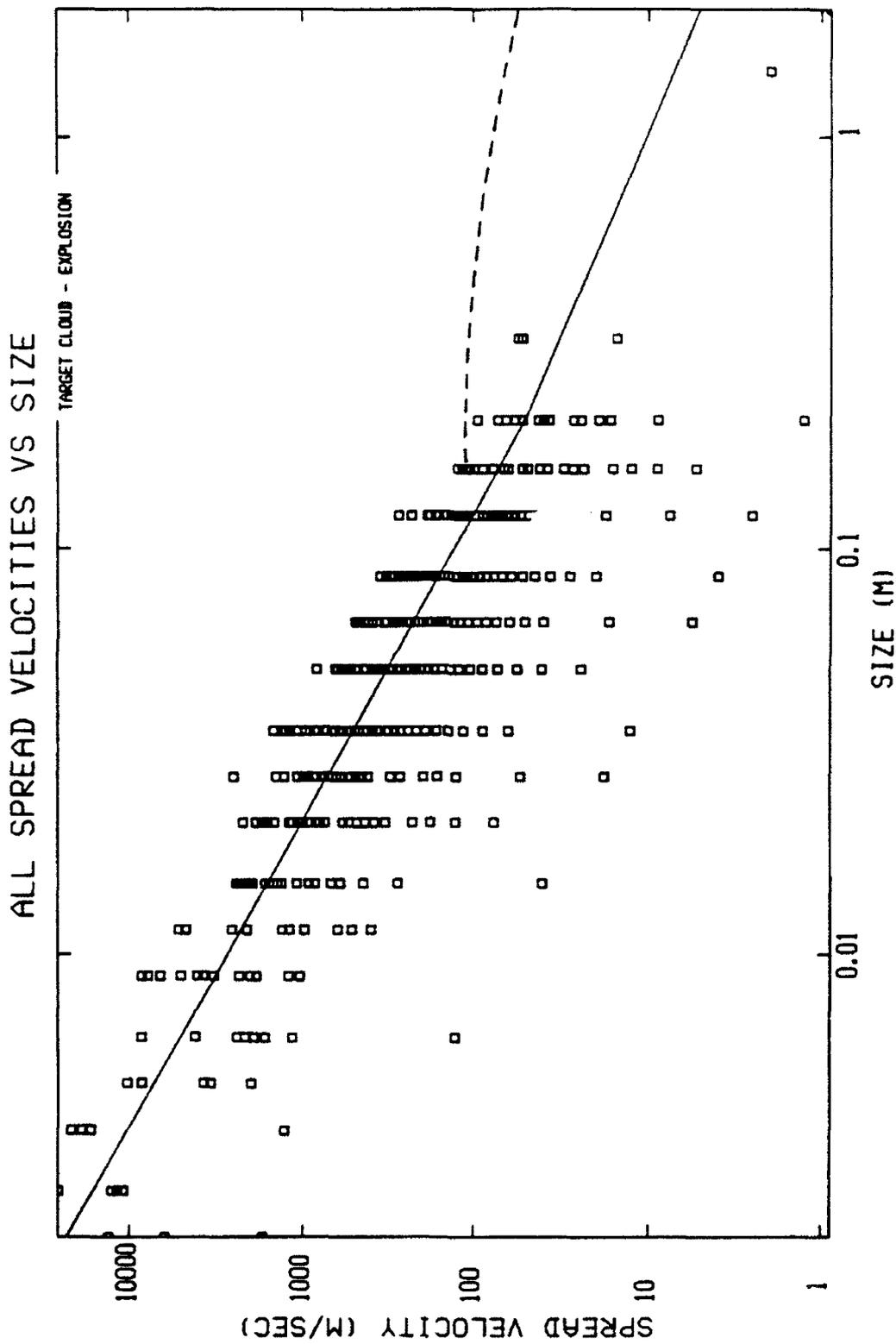


Fig.4. Mass distribution and cumulative mass distribution of fragments of hypothetical breakup event of NOAA-3 Rocket by impact with projectile of mass 2.5 kg at relative velocity of 7 km/s.

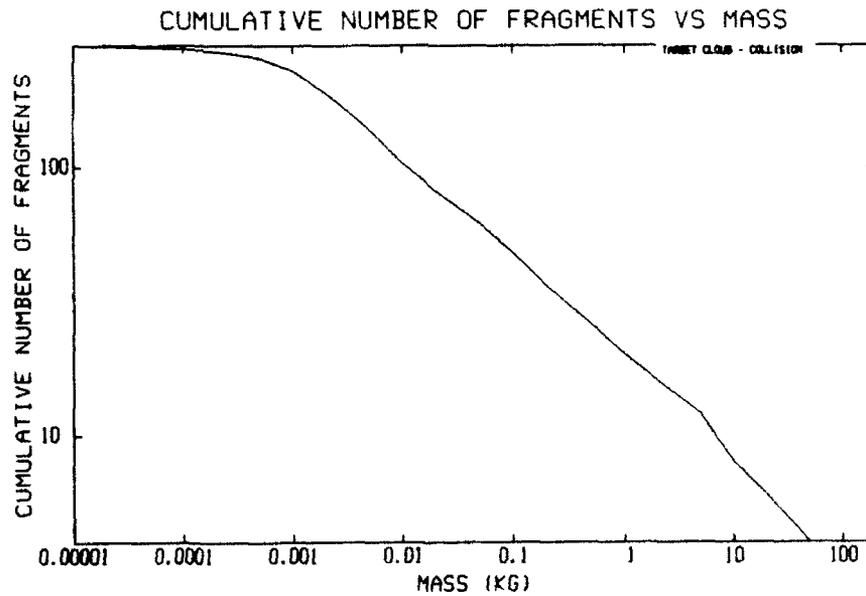
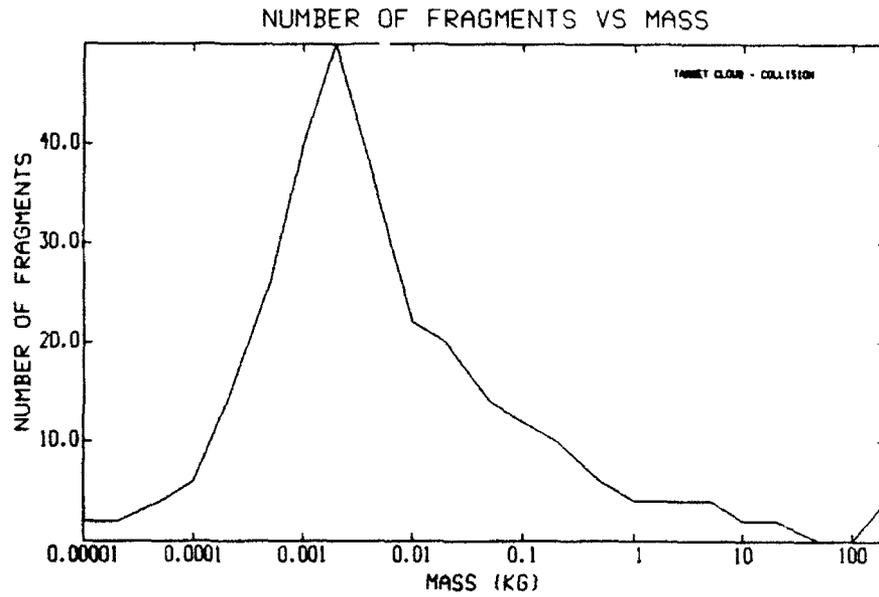


Fig.5. Hypervelocity collision of cylindrical drum by small projectile according to 2-d MAGI Code (side view).

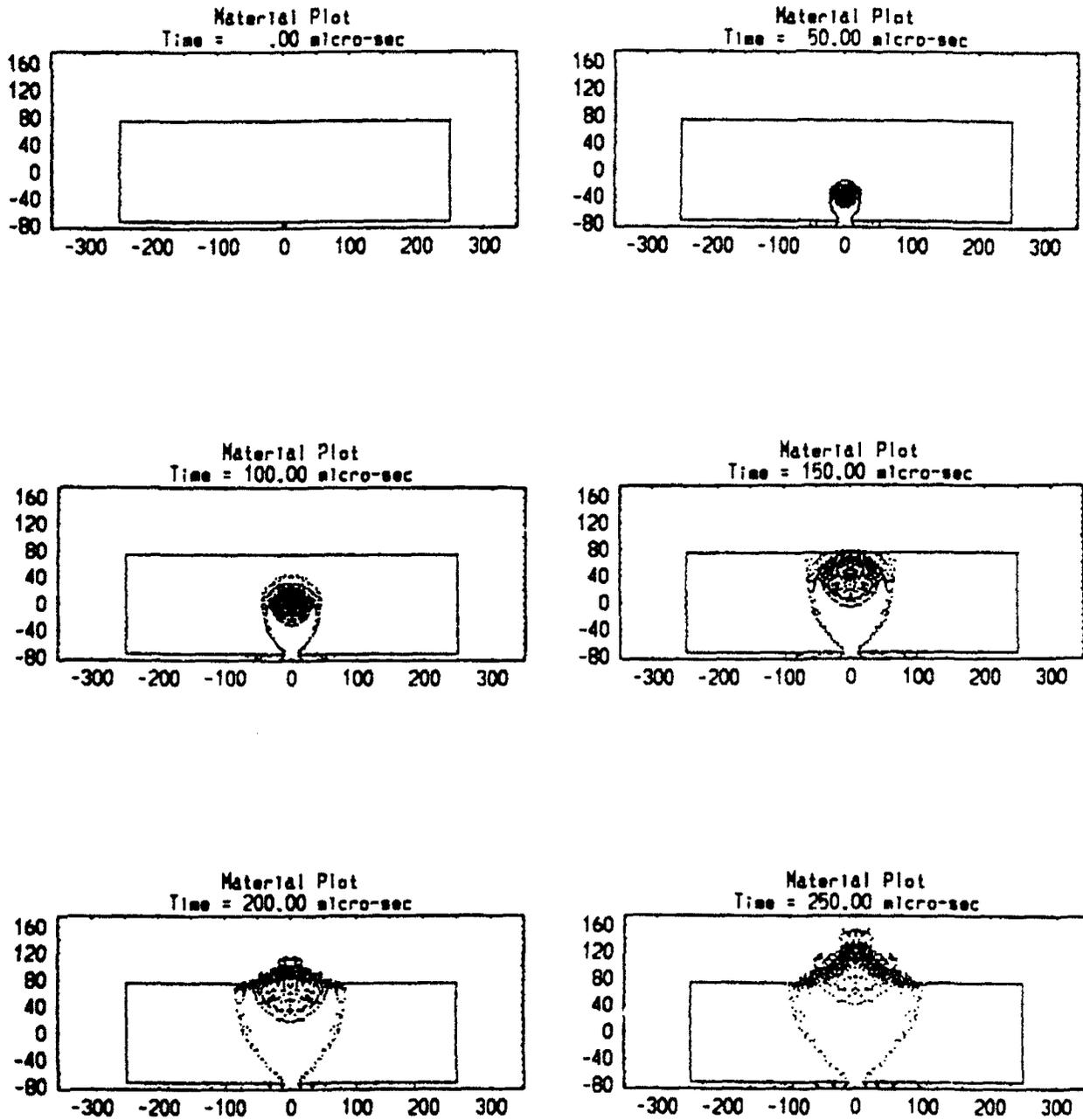
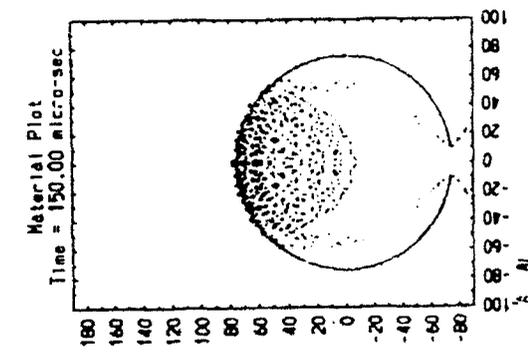
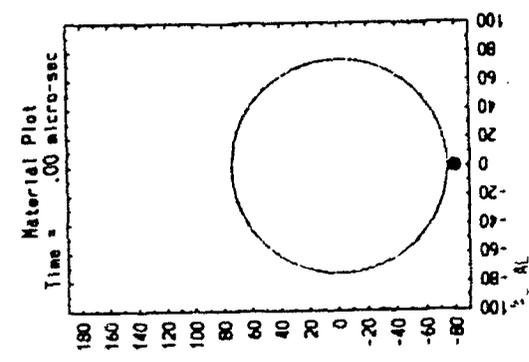
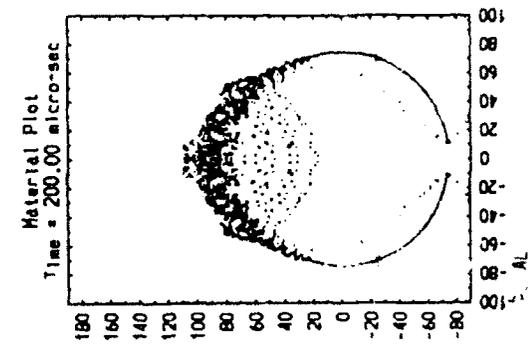
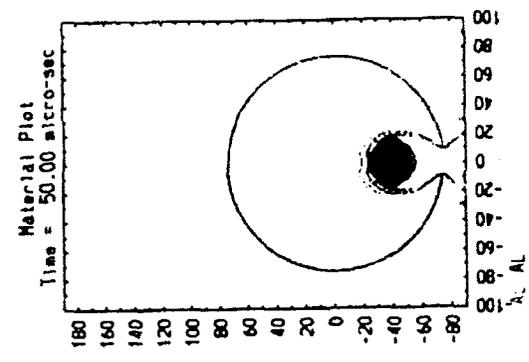
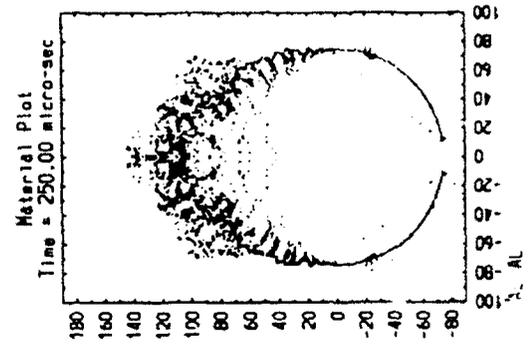
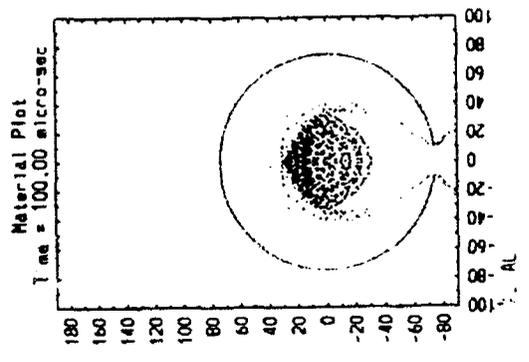


Fig.6. Hypervelocity collision of cylindrical drum by small projectile according to 2-d MAGI Code (axial view).



SPECTROSCOPIC DATA OF ATMOSPHERIC INTEREST

R. H. Tipping
Professor
Department of Physics and Astronomy

University of Alabama
Tuscaloosa, AL 35487-0324

Final Report for:
Summer Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

September 1992

SPECTROSCOPIC DATA OF ATMOSPHERIC INTEREST

R. H. Tipping
Professor
Department of Physics and Astronomy
University of Alabama

Abstract

Spectroscopic data, including such important parameters as rotation-vibrational transition frequencies and intensities, have been calculated for three diatomic molecular species: $^{14}\text{N}^{15}\text{N}^+$, $^{16}\text{O}^{18}\text{O}^+$, and OH. Although the heteronuclear nitrogen and oxygen molecular ions have low concentrations in the Earth's atmosphere, they are polar and, consequently, are infrared active. The fundamental transitions were looked for in atmospheric spectra obtained by researchers at the Phillips Laboratory but were not found. In contrast, highly rotationally excited OH pure rotational transitions were recently observed. Because previous intensity calculations were made assuming a constant dipole moment and these become progressively inaccurate as the rotational quantum number increases, we have calculated improved values that should be of use in the analysis of the OH excitation mechanism.

INTRODUCTION

A number of diatomic species, both neutral and singly ionized, play an important role in the upper atmosphere. In particular, these are important not only from the radiative standpoint, but also from the chemical standpoint because they react strongly either with electrons in the case of ions or with other atomic and molecular species in the case of neutral OH. The role of OH as the source of the nightglow has been known for a long time.¹ In fact, the rotational temperature within a given vibrational band (as determined by emission intensities) has been used to infer the altitude dependence of the OH vibrational population.² Recently, however, pure rotational transitions between highly excited levels have been reported from the CIRRIIS 1A experiment;³ these levels have excitation energies as high as 2.5 eV which implies a non-equilibrium formation mechanism.

In order to investigate the origin of these highly excited rotational levels, one must first determine the populations of the levels and effective rotational temperatures from an analysis of the emission intensities. To do this, one needs accurate theoretical intensities. Before the present work was undertaken, the best intensities for the high rotational lines of the pure rotational transitions were those calculated by Goldman et al.⁴ These results were obtained using a constant dipole moment function which is satisfactory for low rotational levels but becomes progressively less accurate as the rotational excitation increases. We have therefore recalculated the intensities of

these transitions using a more realistic dipole moment function⁵ along with numerical wavefunctions⁶ and these results have been accepted for publication.⁷

In the following section, the method used for the calculation of the spectroscopic parameters for three diatomic molecular species: $^{14}\text{N}^{15}\text{N}^+$, $^{16}\text{O}^{18}\text{O}^+$, and OH is described. In the final section, conclusions concerning the results of the present study are given and additional work necessary in order to ascertain the mechanism responsible for the formation of highly excited OH is discussed.

THEORETICAL METHOD

In order to calculate the transition frequencies and intensities of the heteronuclear ions, we first calculated accurate Dunham parameters using corresponding parameters of the homonuclear ions and their isotopic dependence.⁸ These enabled us to calculate the transition frequencies, and using a standard RKR program, to calculate a numerical potential energy curve. This in turn was used as input for a program which calculates numerical wavefunctions. The dipole moment function can be accurately modeled by considering the electronic charge multiplied by the displacement between the geometric center and the center of mass. Accordingly, for $^{14}\text{N}^{15}\text{N}^+$ and $^{16}\text{O}^{18}\text{O}^+$ these can be written in the form:

$$\text{and } M(x) = 9.24 \times 10^{-2} + 9.42 \times 10^{-2} x$$
$$M(x) = 1.58 \times 10^{-1} + 1.58 \times 10^{-1} x$$

where $M(x)$ is in Debye, and $x = (R - R_e)/R_e$ is the dimensionless displacement from equilibrium. Using these and the numerical wavefunctions, we then calculated the dipole moment matrix elements and Einstein A-coefficients for pure rotational and fundamental vibration-rotational transitions.

We then attempted to look for transitions arising from these ions in atmospheric spectra obtained by researchers at the Phillips Laboratory. Although the strengths of these transitions are appreciable, their abundances are very small owing to the small isotope abundances of ^{18}C and ^{15}N . As a consequence, we were unable to identify absorption features arising from these species. We note, however, that during strong auroral activity, the concentrations of N_2^+ and O_2^+ increase dramatically and as a result transitions arising from the heteronuclear counterparts may become observable. If this were the case and one could identify these transitions and measure their intensity, one could thereby determine the abundances of these molecular ions.

For OH similar calculations were carried out. The frequencies in this case were generated for the 2π ground state by diagonalizing the Hamiltonian as described in the paper by Goldman et al.⁹ As mentioned in the Introduction, the intensities for the pure rotational transitions were then calculated using a realistic dipole moment function⁵ along with numerically generated wavefunctions.⁶ The details of these calculations are given in Ref. 7, so here we only discuss the results.

We calculated intensities (Einstein A values) for the strongest 1793 pure rotational transitions; of these, 1476 agreed with the previous results to within $\pm 5\%$. The remaining 317 A values were distributed according to the results listed in Table 1. As expected, the discrepancies become progressively larger as the vibrational and rotational quantum numbers increase. The largest discrepancy found was 46% for the 4-4 $R_{R22}(48.5)$ pair of lines at 971.8 cm^{-1} .

DISCUSSION AND CONCLUSIONS

As described above, we have calculated transition frequencies and Einstein A values for a large number of vibration-rotational transitions for $^{14}\text{N}^{15}\text{N}^+$, $^{16}\text{O}^{18}\text{O}^+$ and OH. Atmospheric spectra were examined to see if the molecular ions lines were present but they were not found. On the other hand, recent observations of highly excited pure rotational lines of OH should provide some important information concerning the formation and excitation mechanisms. The improved line strengths calculated in the present work will be useful in this connection. In addition, a recalculation of the frequencies and strengths of the vibration-rotational lines is also warranted.

The same theoretical method and computer programs used in the present work can also be used to calculate similar data for other 2p state molecules of atmospheric interest; for instance, NO or C θ O. Indeed, recent measurements of branching ratios for infrared vibrational emission from NO highlight the importance of

having accurate theoretical A values for this species.¹⁰ Using the more accurate frequency data¹¹ and dipole moment function, one should be able to improve the theoretical values calculated a number of years ago by Billingsley.¹² In order to calculate high vibrational transitions, the dipole moment function has to be accurately modeled over a large range of internuclear separations. While one can deduce the magnitude of the dipole moment function and its first few derivatives at the equilibrium internuclear separation from laboratory intensity measurements,¹³ one cannot use a series representation because extrapolation to large R diverges. However, one can combine the near-equilibrium data with accurate ab initio calculations valid at large separations by writing the dipole moment function in the form of a Pade approximant that has the correct long-range behavior and limit.¹⁴ This should provide an accurate representation over the range of interest and allow one to calculate accurate intensities.

One final point is worth noting. The accurate numerical wavefunctions that one can generate can be used for other calculations. For instance, they can be used to calculate scattering cross-sections for vibrational and/or rotational excitation using theoretical formulations such as the impulse approach.¹⁵ These types of calculations are important if one wants to understand the excitation and collisional deactivation of OH molecules in the upper atmosphere. Work along these lines is in progress.

ACKNOWLEDGMENTS

The author would like to thank Dr. W. Blumberg of the Phillips Laboratory (GPOS) for the hospitality during the time when this work was carried out. Also, I would like to thank Dr. H. Dothe for help with the computer calculations and Drs. R. D. Sharma and D. R. Smith for helpful discussions.

REFERENCES

1. A. B. Meinel, AP. J. 111, 555 (1950).
2. G. G. Sivjee and R. M. Hamwey, J. Geophys. Res. 92, 4663 (1987).
3. D. R. Smith, W. A. M. Blumberg, R. M. Nadile, S. J. Lipson, E. R. Huppi, and N. Wheeler, Geophys. Res. Let. 19, 593 (1992).
4. A. Goldman, D. G. Murcray, D. L. Lambert and J. F. Dominy, Mon. Not. R. Astr. Soc. 203, 767 (1983).
5. D. D. Nelson, A. Schiffman, D. J. Nesbit, J. J. Orlando and J. B. Burkholder, J. Chem. Phys. 93, 7003 (1990).
6. D. Goorvitch and D. C. Gallant, JQSRT 47, 505 (1992).
7. D. Goorvitch, A. Goldman, H. Dothe, R. H. Tipping and C. Chackerian, Jr., J. Geophys. Res., to be published (1992).
8. K. P. Huber and G. Herzberg, Constants of Diatomic Molecules, Van Nostrand Reinhold, New York, 1979.
9. A. Goldman, J. R. Gillis and J. A. Coxon, JQSRT 29, 469 (1983).
10. W. T. Rawlins, M. E. Fraser, S. M. Miller and W. A. M. Blumberg, J. Chem. Phys. 96, 7555 (1992).
11. A. H. Saleck, G. Winnewisser and K. M. T. Yamada, Mol. Phys. 76, 1443 (1992).
12. F. P. Billingsley, J. Mol. Spectrosc. 61, 53 (1976).
13. A. S. Pine, A. G. Maki and N.-Y. Chou, J. Mol. Spectrosc. 114, 132 (1985).
14. R. H. Tipping and C. Chackerian, Jr., J. Mol. Spectrosc. 88, 352 (1981).
15. R. D. Sharma, P. M. Bakshi and J. M. Sindoni, Phys. Rev. A 43, 189 (1991).

Table 1. Distribution of A values for the pure rotational transitions of OH having discrepancies greater than $\pm 5\%$ from previous calculations;⁴ v'' and J'' are the ground state vibrational and rotational quantum numbers, respectively.

v''	Number	Type	J''
0	20	R_R	>39.5
	23	S_R	>15.5
	24	R_Q	>26.5
1	4	S_R	>19.5
	13	R_Q	>27.5
2	22	R_R	>43.5
	5	S_R	>25.5
	6	R_Q	>40.5
3	50	R_R	>36.5
	9	S_R	>21.5
	15	R_Q	>32.5
4	82	R_R	>28.5
	16	S_R	>15.5
	28	R_Q	>21.5

THE ROLE OF ATOMIZATION IN LIQUID PROPELLANT ROCKET
COMBUSTION INSTABILITY

Robert J. Turnbull
Professor
Department of Electrical and Computer Engineering
University of Illinois
1406 West Green Street
Urbana, IL 61801

Final Report for:
Summer Research Program
Phillips Laboratory
Edwards Air Force Base

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D. C.

September 1992

THE ROLE OF ATOMIZATION IN LIQUID PROPELLANT ROCKET COMBUSTION INSTABILITY

Robert J. Turnbull
Professor
Department of Electrical and Computer Engineering
University of Illinois

Abstract

A study was made to determine fruitful areas for research on combustion instabilities in liquid fuel rockets. These instabilities involve the interaction of acoustic waves with fuel combustion processes. Atomization of the propellants, specifically the final droplet size, is one of the crucial parameters that determine the burning rate. The size of the droplets resulting from secondary atomization is susceptible to change by acoustical disturbances, thus it could couple into an instability. An investigation was begun on the effects of droplet shape, droplet orientation, and sound wave frequency on the interactions between drops and acoustic waves. The goal is to determine how the secondary atomization process is affected by sound waves and, in turn, how this interaction can produce an instability. Finally possible instability control techniques were examined. Further work in these areas is planned.

THE ROLE OF ATOMIZATION IN LIQUID PROPELLANT ROCKET COMBUSTION INSTABILITY

Robert J Turnbull

INTRODUCTION

Combustion instabilities whose effects range from oscillations in the thrust to the destruction of the rocket occur in liquid fuel rockets. Two distinct types of instabilities have been identified. Interactions of the liquid feed system with the combustion process can cause low frequency or chug instabilities and interactions of the acoustic modes of the combustion chamber with the combustion process can cause medium and high frequency instabilities. Disturbances, either infinitesimal or finite, in the combustion process trigger the instabilities which then cause the disturbances to grow. The growth rate and the ultimate amplitude of the instability determines the possible damage. The research reported on here focuses on the high frequency instabilities which are growing acoustic modes.

An acoustic wave in a combustion chamber experiences some damping so that energy must be added to the wave in order to produce an instability. Important damping mechanisms include viscous losses in liquid drops set into oscillation by the wave and energy loss out the throat of the nozzle. In order for an instability to occur the losses must be offset by a larger energy gain from the combustion process. The energy gain must be at the same frequency as the wave and must have the correct phase relationship with the wave for the wave to grow. These restrictions on the energy addition effectively require that the sound wave itself produce a time-varying change in the combustion rate. Two candidate processes that the wave can affect are the evaporation of a liquid drop or the atomization process. This work, which concentrates on the atomization process, has as its main purpose the design of a research program to help understand and control combustion instabilities. The report starts with a literature review, continues with some

background information and the results of some preliminary calculations and concludes with suggestions for further research.

LITERATURE REVIEW

Combustion Instabilities

"Liquid Propellant Rocket Combustion Instability"¹ is the most valuable and complete reference in the field. It contains not only a discussion of the instabilities but also background material in such areas as fuel injection, atomization, vaporization and combustion. Its main problem is that it is twenty years old. In recent years a number of workshops on combustion instabilities in liquid fuel rockets have been held.^{2,3,4,5} The reports of these workshops are valuable sources for general information on the status of the field. Work in progress designed to understand the fundamental mechanisms of high-frequency instability in liquid rocket engines has been reported.^{6,7} A recent review of the mechanisms of rocket engine combustion instability found that in certain circumstances that any of the processes of injection, atomization, vaporization, mixing, and chemical kinetics can be the dominant process in the instability.⁸

Numerical modeling of the combustion process in liquid rockets is currently an active area of research.⁹ A code has been constructed and validation data is currently being compiled.¹⁰ Modeling has been done on the dynamic response of the feed system¹¹ and a code has been applied to study the combustion process for the Variable Thrust Engine for the Orbit Maneuvering Vehicle.¹²

A theoretical investigation of the suppression of longitudinal combustion instability by active control was performed.¹³ Conditions for the suppression of the instability through pressure transducers were developed assuming that the system parameters were known. Active control has actually been used to control an instability in a turbulent combustor.¹⁴ Again the control was through the use of pressure transducers.

Atomization and Droplet Instabilities

One of the most important determinants of the performance and stability of liquid fuel rocket engines is the atomization of the liquid fuel. This process starts with jets of liquid which break up into drops. The drops in turn often break up into smaller drops before vaporization and combustion occur. The liquid jet breakup into drops is due primarily to the difference in velocity between the jets and the gas in the combustion chamber. The phenomenon is similar to the classical Kelvin-Helmholtz instability on flat sheets.¹⁵ A number of experiments have been performed measuring the drop size distribution formed by the breakup of high speed jets under various conditions.^{16,17,18} A complete theoretical prediction of the breakup phenomena is not current possible but partial explanations are possible.¹⁹ A numerical code has been developed which predicts the observed jet breakup phenomena.²⁰ The jet breakup is well illustrated by photographs in *An Album of Fluid Motion*.²¹ The structure of the sprays resulting from jet breakup has been the subject of a number of studies.^{22,23,24}

A liquid drop with a large velocity relative to its surrounding gas will become unstable and break up into smaller drops. Experimental studies of this phenomenon have mostly been conducted using shock tubes to provide the gas flow.^{25,26,27,28} A criterion for the stability of a drop accelerated by a gas flow has been derived and a critical Bond number has been found.²⁹ Other derivations of breakup criteria have been given and applied to experiments.³⁰ A key factor prior to the breakup of a drop is its shape. The shape of a drop under various flow conditions is known.³¹ Also very important in drop shattering are drop oscillations and the results of an investigation of large amplitude oscillations have also appeared in the literature.³²

The specific problems associated with liquid atomization in rockets have been the subject of a review article.³³ Experimental studies of coaxial injectors are current being performed to provide quantitative information about the spray characteristics.^{34,35,36} Other studies to determine the physical processes in injection and atomization of liquid

rocket propellants have been performed.³⁷ Finally, a recent workshop on atomization and sprays reviewed the current state of the field and suggested directions for future research.^{38,39,40}

Vaporization and Combustion

Since these areas are not as central to the research, only a few references will be given. A review of the status of knowledge on liquid and droplet combustion was written by Faeth.⁴¹ Vaporization of drops under super-critical conditions in rocket engines has been considered.^{42,43} An experimental study of the effect of acoustic disturbances on droplet vaporization rates is currently underway.⁴⁴ This should provide information about one way energy can be coupled into combustion instabilities.

BACKGROUND

High frequency combustion instabilities in liquid fuel rockets are acoustic modes whose amplitudes increase in time. A combustion chamber acts as a resonant cavity for sound waves. This cavity, which is usually approximately cylindrical in shape, supports certain acoustic modes. That is, only certain discrete frequencies can exist in the chamber.⁴⁵ A growing acoustic wave must have a growing energy content. Since both energy loss and energy gain processes are present, the gain processes must be larger than the loss processes. Among the energy loss mechanisms are a loss of energy out the nozzle throat and viscous losses in the wave. The presence of drops of liquid propellant greatly increases the losses since the sound waves set up oscillations in the drops and the viscosity of liquids is usually much larger than that of gases. Another source of losses can be produced by adding passive damping structures to the combustion chamber for that purpose. The energy gain in the waves must come from the energy released by combustion. If energy is to be added to the wave by the combustion process, the combustion process must have a time-varying component at the same frequency as the

acoustic wave. The other criterion is that the extra combustion energy release must be in phase with the wave. For example, a sinusoidal sound wave has a pressure which varies sinusoidally in time. If a combustion process is taking place that is affected by the wave the pressure due to the combustion also has a sinusoidal component at the same frequency as the wave. If the phase difference between the peak pressure in the wave and the peak pressure due to the combustion is less than 90° , then energy will be added to the wave. On the other hand, if the phase difference is greater than 90° , the combustion process will actually damp the wave. Therefore an instability requires a combustion process with a time-varying component of the correct frequency and phase.

In order for the atomization process to be the driving process for the combustion instability, the sound wave must affect the droplet size. The sound wave consists of time-varying pressure and velocity. Because the droplet size is small compared to the sound wavelength, the pressure variations experienced by a drop will be spatially uniform. This will just raise and lower the pressure in the drop but have no effect on the droplet size. On the other hand the gas velocity produced by the wave will produce a non-uniform pressure on the drop and therefore it has the potential of affecting the atomization process. Three possible ways that the sound wave produced gas velocity could affect the atomization process are: 1) The wave could produce high enough velocities to shatter all drops above a certain size. 2) The wave could couple with droplet oscillations to shatter the drops. 3) Highly elongated drops resulting from the breakup of the jets could be shattered by the wave. These three categories are not mutually exclusive. The assumption is all this is that the primary atomization, that is, the initial breakup of the liquid jet is due just to the relative velocity between the jet and the gas in the combustion chamber and is not appreciably affected by the sound wave.

The first breakup mechanism, drops of arbitrary shape being shattered by the sound wave driven gas velocity, is the phenomena reported in the literature discussed above. It requires a large amplitude wave. (The experiments were done using shock

waves.) This could be the mechanism for combustion instabilities which require a finite amplitude wave for initiation. However, it could cause instabilities which grow from small disturbances.

If a sound wave is to couple with drop oscillations, the sound wave frequency (or a harmonic for non-sinusoidal waves), must be approximately equal to the natural frequency for droplet oscillations. Since the non-uniform pressure in the drop is proportional to the velocity squared and the sound wave produced velocity is superposed on the relative velocity between the drop and the gas, the non-uniform pressure in the drop could have components at both the sound wave frequency and twice that frequency. If the sound wave is propagating parallel to the gas velocity, i.e., if we have a longitudinal wave, the frequency of excitation of the drop oscillations is the sound wave frequency. However, if there is a transverse wave, the droplet excitation is at twice the sound wave frequency. A problem with this mechanism is that the drops can only undergo a few oscillations before they evaporate, so the coupling must be strong.

Droplets resulting from the primary atomization of a jet are often formed with a highly non-spherical shape. Figure 1 shows an example of this. These elongated drops would be very susceptible to breakup by a sound wave of small or moderate amplitude. The orientation of the drops with respect to the wave is extremely important. If the sound wave frequency was near the natural frequency for drop oscillations, the effect would be greatly enhanced.



Figure 1 Photograph of the atomization of a high-speed jet.⁴⁶

The feedback process by which atomization can provide energy to produce an instability is the following. Small drops evaporate more rapidly than large drops and therefore combustion of small drops is faster than that of large drops. If the sound wave affects the secondary atomization process, the average drop size will vary in time with the same frequency as the sound wave. The energy released by combustion will then have a time-varying component at the same frequency as the sound wave and this energy release will couple to the sound wave. Because there is a time-delay between atomization and combustion, no definitive statement can be made about the phase between the wave and the energy release. At least for some wave frequencies, instability is a definite possibility.

EFFECTS OF AN ACOUSTIC WAVE ON A DROP

Some calculations were done to illustrate the effect of a sound wave on a drop. Since it is the gas velocity in the wave which could cause a drop to shatter, the calculations were of the effect on a drop of a velocity in the gas surrounding the drop. The calculations were of the forces on a non-spherical drop and the effects of a time varying force.

Forces on a non-spherical drop due to flow around the drop

The non-spherical drops were approximated as prolate and oblate spheroids. If the velocity field is approximately that of an inviscid fluid, analytical expressions can be obtained both for the velocity and the pressure. Calculations were done for various values of oblateness and prolateness and for velocities along each of the principal axes of the spheroids. The results for oblate spheroids are shown in Figures 2 and 3. The pressure differences are normalized to $\frac{1}{2}\rho V_0^2$ where V_0 is the velocity far from the drop. The horizontal axis of the plots represent the ratio of the major and minor axes of the spheroid so that a value of 1 is a sphere. It is apparent from these figures that a

velocity parallel to the minor axis has a much larger effect on the drop. The shape effects are also more dramatic for velocities in this direction. A velocity in the direction of the minor axis exerts forces which tend to increase the oblateness of the spheroid.

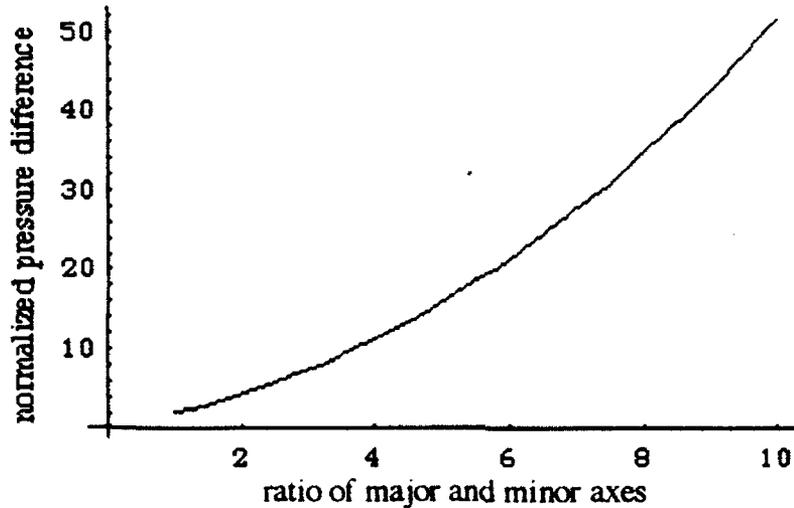


Figure 2. Difference in pressure between the poles and the equator for an oblate spheroid with a velocity parallel to the minor axis.

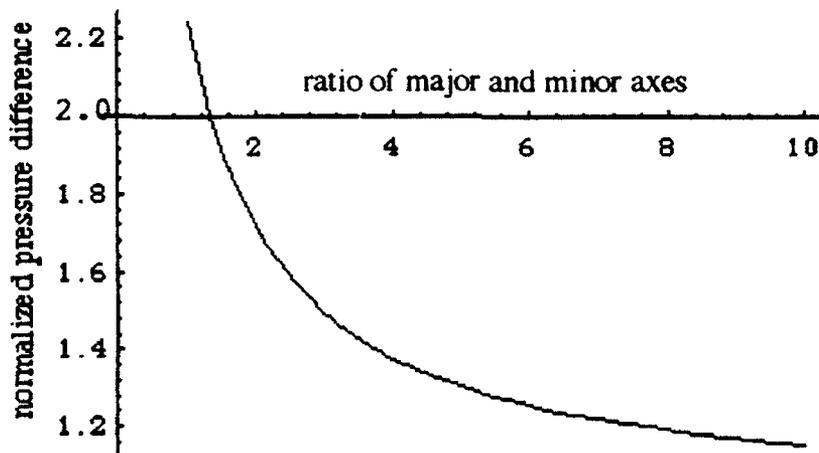


Figure 3. Difference in pressure between the poles and the equator for an oblate spheroid with a velocity perpendicular to the minor axis.

The prolate spheroid is probably of more interest than the oblate spheroid since it approximates elongated drops which are more susceptible to being shattered. The results of calculations on prolate spheroids are shown in Figures 4 and 5. Figure 4 shows the

effect of velocity parallel to the major axis and Figure 5 shows the effect of flow perpendicular to the major axis. The plot axes are the same as for the oblate spheroid, In the case of a prolate spheroid a sound wave traveling perpendicular to the major axis has more effect. What is not shown here is the effect of surface tension which assist in breaking up an elongated drop.



Figure 4. Difference in pressure between the poles and the equator for an prolate spheroid with a velocity parallel to the major axis.

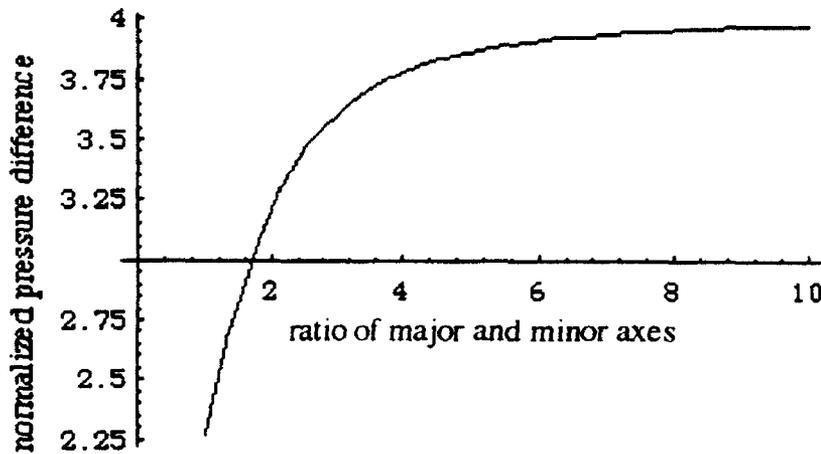


Figure 5. Difference in pressure between the poles and the equator for an prolate spheroid with a velocity perpendicular to the major axis.

Interaction of a sound wave with droplet oscillations

For these calculations a velocity was assumed and the response of the drop to the velocity was calculated. First, so as to provide a benchmark, a spherical drop was assumed to be placed in a steady flow with a Weber number of .1. The resultant drop oscillations are shown in Figure 6. What is illustrated is the time variation in radius at the poles of the drop as a function of time. The radius fluctuations are normalized to the equilibrium radius. The time scale is also normalized and the normalization is the same for all the plots in this section. The drop is assumed to be spherical and motionless at time $t=0$. It can be seen that a new equilibrium is established with the polar radius decreased by about 6%. The drop oscillates about this equilibrium. All equations are linearized so that the oscillations are sinusoidal.

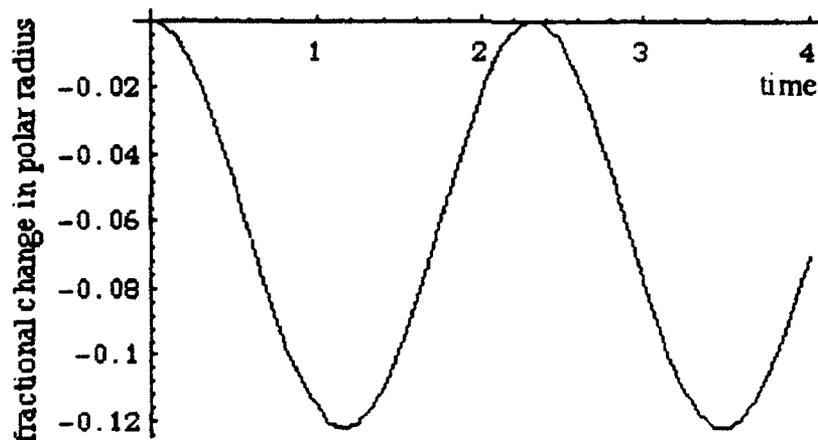


Figure 6. Time variation of the polar radius of a drop in a constant gas flow.

Next, the gas velocity was assumed to be due to a sound wave and therefore varied sinusoidally in time. The results of these calculations are shown in Figures 7, 8, and 9. In Figure 7 the sound wave frequency is approximately equal to the frequency of drop oscillations. In Figure 8 the sound wave frequency is approximately twice the frequency of droplet oscillations while in Figure 9 the sound wave frequency is approximately half the droplet oscillation frequency. The maximum Weber number in each case is .1, the same as in Figure 6.

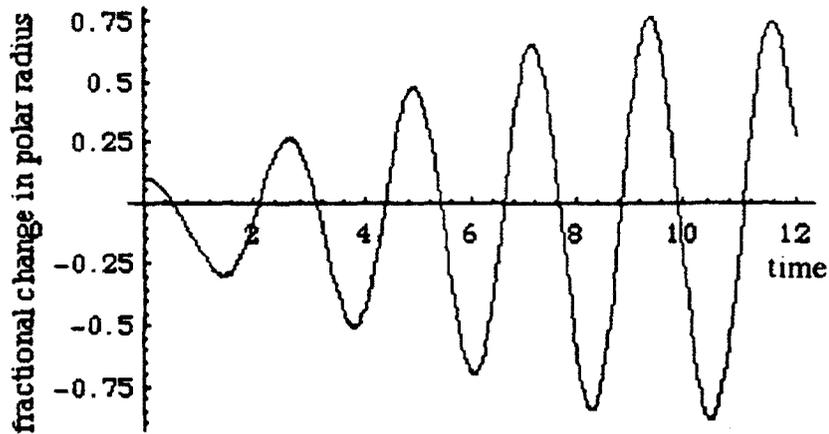


Figure 7. Time variation of the polar radius of a drop in a sound wave whose frequency equals that of the droplet oscillations.

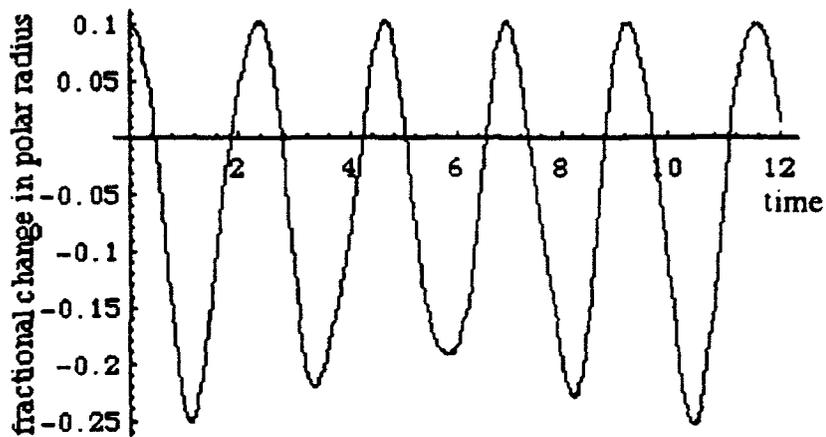


Figure 8. Time variation of the polar radius of a drop in a sound wave whose frequency is twice that of the droplet oscillations.

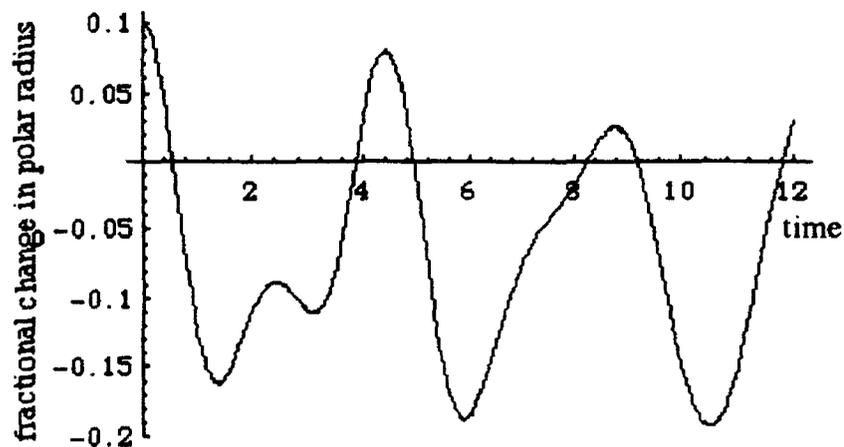


Figure 9. Time variation of the polar radius of a drop in a sound wave whose frequency is half that of the droplet oscillations.

In these figures, the effect of a sound wave at the droplet oscillation frequency can be seen dramatically. In a matter of a few oscillations a velocity whose magnitude was only large enough to produce a negligible drop deformation, can cause a drop to shatter. Because the equations were linearized, shattering is not predicted but oscillations of this amplitude would surely shatter the drop. On the other hand, frequencies very different from the droplet oscillation frequency cause oscillations whose amplitudes are not much greater than those of a steady flow with the same velocity. It should be noted that only the fundamental mode of oscillation was considered. Higher frequency sound waves could couple to higher modes of droplet oscillations.

CONTROL OF INSTABILITIES BY PARAMETER VARIATIONS

In order to get the maximum performance out of a liquid fuel rocket, it might be desirable to operate under conditions conducive to combustion instabilities. An active control system could work by measuring the pressure at various points in the combustion chamber and applying signals to transducers to cancel any acoustical disturbances.^{47,48} However, signals applied to the propellant system would have a greater effect than those applied in the combustion chamber. The signals could change the feed rate or possibly apply a charge to the liquid. In order to be effective, the effects of acoustic waves on atomization must be better understood.

Active control has the disadvantage that instabilities must be detected, the signals must be processed and the proper correction sequence initiated, all in real time. This raises the question of the possibility of using a passive control system instead. A passive system would vary the parameters in a manner that would prevent the instabilities from growing indefinitely. Instabilities involve waves propagating through the combustion chamber numerous times. A control scheme might consist of making sure that a wave that is amplified on one pass through the chamber is damped on the next pass. This

requires a detailed knowledge of the instability mechanisms and it may not be possible even then.

Some simulations of passive control of hypothetical systems were performed. If the controller was assumed to know nothing about the instability, the results were not encouraging. The results from two of the simpler simulations are shown in Figure 10. In these simulations the system was randomly switched between a stable state and an unstable state. (This could correspond to a wave being amplified on one pass and damped on the next pass through the combustion chamber.) The parameters were picked so that, on average the system was neutrally stable. The amplitude is plotted as a function of time with the initial amplitude set to 1. It can be seen that the random parameter variations can produce an instability by themselves. In the first simulation, the amplitude grows to more than 500 times the initial value while in the second simulation the amplitude shrinks to approximately zero. On the average the systems switched from growth to damping every one unit in time. The period of the wave was two time units.

SUGGESTIONS FOR FURTHER WORK

The most important problem for further work is the study of the interaction of acoustic waves with drops. This work should be primarily experimental. Droplets should be generated and subjected to strong sound waves. The sound wave frequencies should range from well below the droplet oscillation frequency to well above it. This will allow the effects of any resonances to be observed. Both spherical and non-spherical drops and all orientations of the major axes of the drops to the direction of sound wave propagation should be studied. One way of producing non-spherical drops is to charge the drops when they are formed and apply a strong electric field to the drop. Varying droplet size may be easier than varying the frequency of the sound waves since the system could then operate at an acoustical resonance.

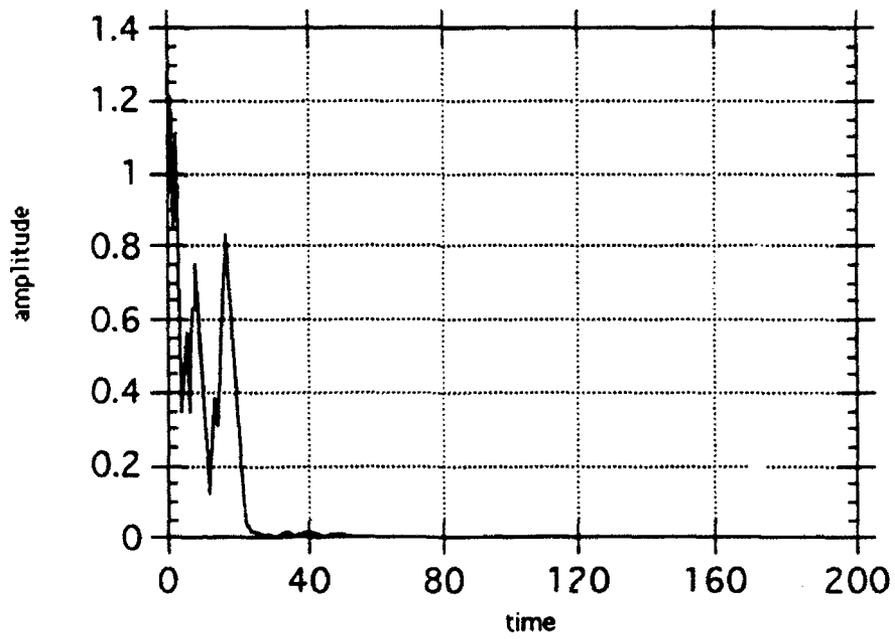
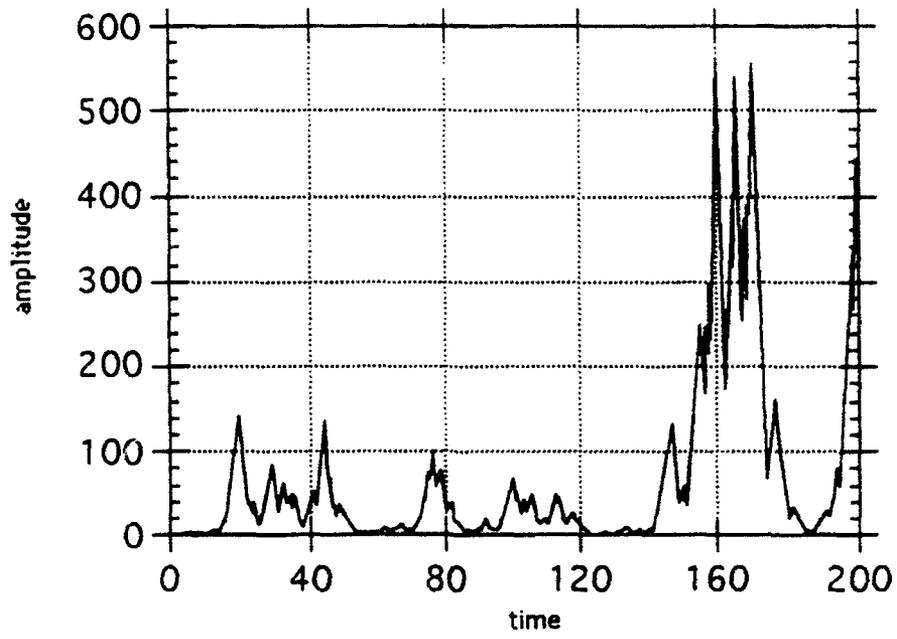


Figure 10. Simulations of the effects of random parameter changes on an wave.

A secondary area for further research is to try to develop control strategies. This could use the results of the drop work. The effect of charging the drops or changing the drop orientation on the growth rate could be studied. Finally, work on the possibility of a passive control strategy should continue.

REFERENCES

- ¹ David T. Harrje and Frederick H. Readon, eds., *Liquid Propellant Rocket Combustion Instability*, NASA technical report SP-194(1972).
- ² Philip A. Kessel and Elizabeth Slimak, "Workshop Report, Unstable Combustion in Liquid Rocket Engines," 25th JANNAF Combustion Meeting, CPIA-Pub 498, 3, 349(1988)
- ³ R. J. Jensen, "A summary of the JANNAF Workshop on Liquid Rocket Engine Combustion Driven Instability Mechanisms," 26th JANNAF Combustion Meeting, CPIA-Pub 529, 2, 273(1989)
- ⁴ Richard E. Walker, "Liquid Rocket Engine Combustion Instability Hardware Design Concerns," 27th JANNAF Combustion Meeting, CPIA-Pub 557, 3, 317(1990).
- ⁵ M. D. Klem, "JANNAF Liquid Rocket Combustion Instability Panel Research Recommendations," 27th JANNAF Combustion Meeting, CPIA-Pub 557, 3, 345(1990).
- ⁶ G. B. Cox, "The Liquid Stability Mechanisms Program," 27th JANNAF Combustion Meeting, CPIA-Pub 557, 3, 263(1990).
- ⁷ G. B. Cox and E. L. Petersen, "Liquid Stability Mechanisms Program Summary," 28th JANNAF Combustion Meeting, CPIA-Pub 573, 2, 51(1991).
- ⁸ W. E. Anderson, H. M. Ryan, and R. J. Santoro, "Combustion Instability Phenomena of Importance to Liquid Bi-Propellant Rocket Engines," 28th JANNAF Combustion Meeting, CPIA-Pub 573, 2, 99(1991).
- ⁹ G. B. Cox, "JANNAF Combustion Instability Panel Workshop Report: Numerical Methods in Combustion Instability," 27th JANNAF Combustion Meeting, CPIA-Pub 557, 3, 307(1990).
- ¹⁰ T. J. Nguyen, R. E. Walker, and M. D. Klem, "Test Program to Provide Validation Data for the Rocket Combustor Interactive Design (ROCCID) Code," 28th JANNAF Combustion Meeting, CPIA-Pub 573, 2, 31(1991).
- ¹¹ S. Farhangi, "Combined Effect of Fuel/Oxidizer Injection System on Rocket Engine Combustion Stability," 28th JANNAF Combustion Meeting, CPIA-Pub 573, 2, 123(1991).

- 12 A. Krisnan, A. J. Przekwas, and K. W. Gross, "Spray Combustion in Liquid Rocket Engine Thrust Chambers," 28th JANNAF Combustion Meeting, CPIA-Pub 573, 3, 307(1991).
- 13 V. Yang, A. Sinha, and Y.-T. Fung, "State-Feedback Control of Longitudinal Combustion Instabilities," *J. Propulsion*, 8, 66(1990).
- 14 G. Billoud, M. A. Galland, C. H. Huu, and S. Candel, "Adaptive Active Control of Combustion Instabilities," *Combustion Science and Technology*, 81, 257(1992).
- 15 S. Chandrasekhar, *Hydrodynamic and Hydromagnetic Instabilities*, London, Oxford University Press, Chapter 11 (1961).
- 16 R. D. Reitz and F. V. Bracco, "Mechanism of Atomization of a Liquid Jet," *Physics of Fluids*, 25, 1730(1982).
- 17 P.-K. Wu, G. A. Ruff, and G. M. Faeth, "Primary Breakup in Liquid-Gas Mixing Layers," *Atomization and Sprays*, 1, 421(1991).
- 18 G. A. Ruff, P.-K. Wu, L. P. Bernal, and G. M. Faeth, "Continuous- and Dispersed-Phase Structure of Dense Nonevaporating Pressure-Atomized Sprays," *Journal of Propulsion and Power*, 8, 280(1992).
- 19 H. Q. Yang, "Asymmetric Instability of a Liquid Jet," *Physics of Fluids A*, 4, 681(1992).
- 20 P. Y. Liang and R. J. Ungewitter, "Direct Numerical Simulation of Atomization and Free Surface Instabilities from First Principles," 27th JANNAF Combustion Meeting, CPIA-Pub 557, 3, 239(1990).
- 21 M. Van Dyke, *An Album of Fluid Motion*, Stanford, Ca, Parabolic Press, (1982).
- 22 G. A. Ruff, A. D. Sagar, and G. M. Faeth, "Structure and Mixing Properties of Pressure-Atomized Sprays," *AIAA Journal*, 27, 901(1989).
- 23 G. M. Faeth, "Structure and Atomization Properties of Dense Turbulent Sprays," *23rd Symposium on Combustion*, The Combustion Institute, 1345(1990).
- 24 G. A. Ruff, L. P. Bernal, and G. M. Faeth, "Structure of the Near-Injector Region of Nonevaporating Pressure-Atomized Sprays," *Journal of Propulsion*, 7, 221(1991).
- 25 A. A. Ranger and J. A. Nicholls, "Aerodynamic Shattering of Liquid Drops," *AIAA Journal*, 7, 285(1969).
- 26 P. D. Patel and T. G. Theofanous, "Hydrodynamic Fragmentation of Drops," *Journal of Fluid Mechanics*, 103, 207(1981).
- 27 A. Wierzba and K. Takayama, "Experimental Investigation of the Aerodynamic Breakup of Liquid Drops," *AIAA Journal*, 26, 1329(1988).

- ²⁸ T. Yoshida and K. Takayama, "Interaction of Liquid Droplets with Planar Shock Waves," *Journal of Fluids Engineering*, 112, 481(1990).
- ²⁹ E. Y. Harper, G. W. Grube, and I.-D. Chang, "On the Breakup of Accelerating Drops," *Journal of Fluid Mechanics*, 52, 565(1972).
- ³⁰ J. Kitscha and G. Kocamustafaogullari, "Breakup Criteria for Fluid Particles," *International Journal of Multiphase Flow*, 15, 573(1989).
- ³¹ J. R. Grace and M. E. Weber, "Hydrodynamics of Drops and Bubbles," in G. Hetsroni, editor, *Handbook of Multiphase Systems*, New York, McGraw-Hill, 1-204(1982).
- ³² E. Becker, W. J. Hiller, and T. A. Kowalewski, "Experimental and Theoretical Investigation of Large-Amplitude Oscillations of Liquid Droplets," *Journal of Fluid Mechanics*, 231, 189(1991).
- ³³ A. Ghafourian, A. Mahalingham, H. Dindi, and J. W. Daily, "A Review of Atomization in Liquid Rocket Engines," 27th JANNAF Combustion Meeting, CPIA-Pub 557, 3, 253(1990).
- ³⁴ M. Zaller, "Lox/Hydrogen Coaxial Injector Atomization Test Program," 27th JANNAF Combustion Meeting, CPIA-Pub 557, 3, 229(1990).
- ³⁵ M. Zaller, and M. D. Klem, "Coaxial Injector Spray Characterization Using Water/Air as Simulants," 28th JANNAF Combustion Meeting, CPIA-Pub 573, 2, 151(1991).
- ³⁶ S. V. Sankar, a. Brena de la Rosa, A. Isakovic, and W. D. Bachalo, "Experimental Investigation of Rocket Injector Atomization," 28th JANNAF Combustion Meeting, CPIA-Pub 573, 2, 187(1991).
- ³⁷ " N. Chigier, H. Eroglu, U. Shavit, and M. Xu, "Physical Processes Involved in Injection and Atomization of Liquid Propellants," 28th JANNAF Combustion Meeting, CPIA-Pub 573, 2, 175(1991).
- ³⁸ N. Chigier, "Future Research in Atomization and Sprays," in N. Chigier, editor, *Atomization and Sprays 2000*, 1(1991).
- ³⁹ G. M. Faeth, "Structure, Breakup and Turbulence Interaction in Sprays," in N. Chigier, editor, *Atomization and Sprays 2000*, 27(1991).
- ⁴⁰ C. F. Edwards, "Research Directions in Spray Combustion for the Next Few Decades," in N. Chigier, editor, *Atomization and Sprays 2000*, 41(1991).
- ⁴¹ G. M. Faeth, "Current Status of Droplet and Liquid Combustion," *Progress in Energy and Combustion Science*, 3, 191(1977).
- ⁴² Richard J. Priem, "Calculating Vaporization rates in Liquid Rocket Engines Above Critical Pressure-Temperature Conditions," 27th JANNAF Combustion Meeting, CPIA-Pub 557, 3, 195(1990).

- ⁴³ W. A. Sirignano, C. H. Chiang, and J.-P. Delplanque, "Numerical Simulation and Modeling of Lox Droplet Vaporization at Supercritical Conditions," 28th JANNAF Combustion Meeting, CPIA-Pub 573, 2, 69(1991).
- ⁴⁴ M. Winter and T. J. Anderson, "Measurement of the Effect of Acoustic Disturbances on Droplet Vaporization Rates," 28th JANNAF Combustion Meeting, CPIA-Pub 573, 2, 79(1991).
- ⁴⁵ J. W. S. Rayleigh, *The Theory of Sound*, New York, Dover Publications, (1945).
- ⁴⁶ M. Van Dyke, *An Album of Fluid Motion*, Stanford, Ca, Parabolic Press, 105(1982).
- ⁴⁷ V. Yang, A. Sinha, and Y.-T. Fung, "State-Feedback Control of Longitudinal Combustion Instabilities," *J. Propulsion*, 8, 66(1990).
- ⁴⁸ G. Billoud, M. A. Galland, C. H. Huu, and S. Candel, "Adaptive Active Control of Combustion Instabilities", *Combustion Science and Technology*, 81, 257(1992).

EXPERIMENTAL INVESTIGATION OF HOMOGENEOUS AND HETEROGENEOUS
NUCLEATION/CONDENSATION PROCESSES AND PRODUCTS IN COIL.

Philip D. Whitefield
Research Associate Professor
Department of Chemistry

and

W. Mark Barnett
Graduate Student
Department of Chemistry

Cloud and Aerosol Sciences Laboratory,
University of Missouri - Rolla. (UMR/CASL).
Norwood Hall G11
Rolla, MO 65401

Final Report for:
Summer Research Program
Phillips Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, WASHINGTON, D.C.

August 1992.

EXPERIMENTAL INVESTIGATION OF HOMOGENEOUS AND HETEROGENEOUS
NUCLEATION/CONDENSATION PROCESSES AND PRODUCTS IN COIL.

Philip D. Whitefield
Research Associate Professor
and
W. Mark Barnett
Graduate Student
Department of Chemistry Cloud and Aerosol Sciences Laboratory,
University of Missouri - Rolla.

Abstract

This paper describes the preliminary results of an ongoing study to characterize the nature and sources of sub-micron aerosols in the Phillips Laboratory small scale supersonic COIL device and other COIL devices. Aerosols from both sub- and supersonic flow regimes were sampled and characterized using the University of Missouri-Rolla, Mobile Aerosol Sampling System (MASS). Under all operating conditions where the oxygen generator discs were rotating, significant concentrations of aerosols were detected. Typically these aerosols had peak dry diameters of <0.05 microns and nascent wet diameters of <0.08 microns. Their total number density increased with increasing rotating disc velocity and with the addition of chlorine. A maximum number density of $<3000/cc$ was observed at maximum chlorine flow rates when the initial generator mixture had been heavily depleted (i.e. neutralized with chlorine). Experiments to observe homogeneous nucleation of aerosols in simulated supersonic laser gas flows were unsuccessful. The critical supersaturation spectra for a KOH, KCl, I_2 have been measured and compared to theoretical calculations.

EXPERIMENTAL INVESTIGATION OF HOMOGENEOUS AND HETEROGENEOUS
NUCLEATION/CONDENSATION PROCESSES AND PRODUCTS IN COIL.

Philip D. Whitefield
and
W. Mark Barnett

INTRODUCTION.

This report describes how the UMR Mobile Aerosol Sampling System is being used in conjunction with a small scale COIL and a supersonic nozzle flow system, to verify the existence of sub-micron aerosols in COILs, to characterize these species, and to assist in the validation of models addressing the role of such species in COIL performance.

The presence, nature and properties of aerosols in the reactive flow of COIL devices are not well characterized. Excellent beam quality results and the short residence times associated with supersonic flow conditions permit an estimate of less than 0.1 microns for an upper bound on the typical diameter to be expected for such species. Recent anomalous device performance data have been rationalized through models invoking sub-micron aerosols[1]. However, no direct experimental evidence for the creation and or existence of such species exists.

Recent COIL data obtained during the operation of several independent COIL devices demonstrate anomalous pressure increases (typically 38%) in the laser cavity, during a given run[1 - 3]. These pressure increases are estimated to cause on the order of a factor of two reduction in laser output power. Furthermore, these pressure increases are concomitant with significant increases in water vapor concentration approaching 5 mole% of the total laser gas flow. An analysis, intended to rationalize these data, has been performed [1] and has led to the hypothesis of a condensation shock model. In this model the cavity pressure increase is assumed to result from an instantaneous heat release into the reactive flow creating a shock. This heat is assumed to come from the enthalpy of condensation of water droplets that homogeneously nucleate in the extremely high supersaturations of water vapor that occur in the exit plane of a COIL supersonic nozzle. The calculated heat release and system residence time lead to the prediction that the proposed condensation shock will produce water droplets less than 0.002 micron in diameter with a total number density on the order of $1.0E13$ cm⁻³. The precedents for, and effects of, water condensation in

supersonic nozzles have been studied extensively [4], and this model simulates the actual physical performance data of the laser[1].

Identifying homogeneous nucleation in a laser device such as a COIL is complicated by the production of heterogeneously nucleated aerosols in the two phase (gas/liquid) singlet oxygen generator. A schematic diagram of a small scale COIL device is shown in figure (1). The generators employed in COIL devices typically operate at sub-atmospheric pressures (about 50torr) and rely on a two phase heterogeneous reaction between aqueous basic hydrogen peroxide (BHP) and gaseous chlorine to produce gaseous electronically excited oxygen. The generator products including metal salt aerosols, heterogeneously nucleated water and BHP aerosols are swept through liquid separatorss (in two of the three devices cited) and on to the laser nozzle in helium carrier gas. The efficiency of particulate removal by the cold traps is unknown. Inertial and gravitational considerations assist in the efficient removal of large aerosols i.e. those typically with diameters >10 micron. The motion of smaller aerosols is largely Brownian in nature and their removal in the cold traps is expected to be considerably less efficient since these particles will be expected to follow the laser gas flow streamlines. Iodine is introduced into the reactive flow in its gaseous state, having been pre-vaporized. The presence of sub micron aerosols in the nozzle offer the potential for heterogeneous iodine nucleation and reduced device performance efficiency.

This study if performed in the laser device alone has the added complication of having to distinguish between concomitant homogeneously and heterogeneously nucleated aerosols when the sampling port is in the cavity i.e. supersonic region. To avoid this complication it was decided to simulate the condensation shock phenomena in a supersonic nozzle flow system. This facility was essentially the laser flow system operated with no liquid phase reagents in the generator stage (see figure (1)). In the simulation experiment flows of water vapor laden carrier gas were introduced to the flow system in the subsonic generator region and expanded through the supersonic nozzle under the typical laser operating conditions where the shock had been observed. If the shock was driven by purely a homogeneous mechanism this event would be observed without the complication of concomitant heterogeneously nucleated aerosols being carried in the flow from the generator.

The size range of droplets predicted [1], in the absence of coagulation, falls close to the limit of detectability of even the state of the art UMR facilities employed in this study. At the aerosol number densities predicted [1], however, coagulation will readily occur. The degree of coagulation will depend on the residence time. The coagulation rate for 0.002micron aerosols collecting 0.002micron aerosols is $0.968E-8$ cm³/s. this is the number of scavenging events per collector particle per unit concentration of collectees. For an aerosol concentration of $1.0E13$ cm⁻³, the rate of scavenging events per

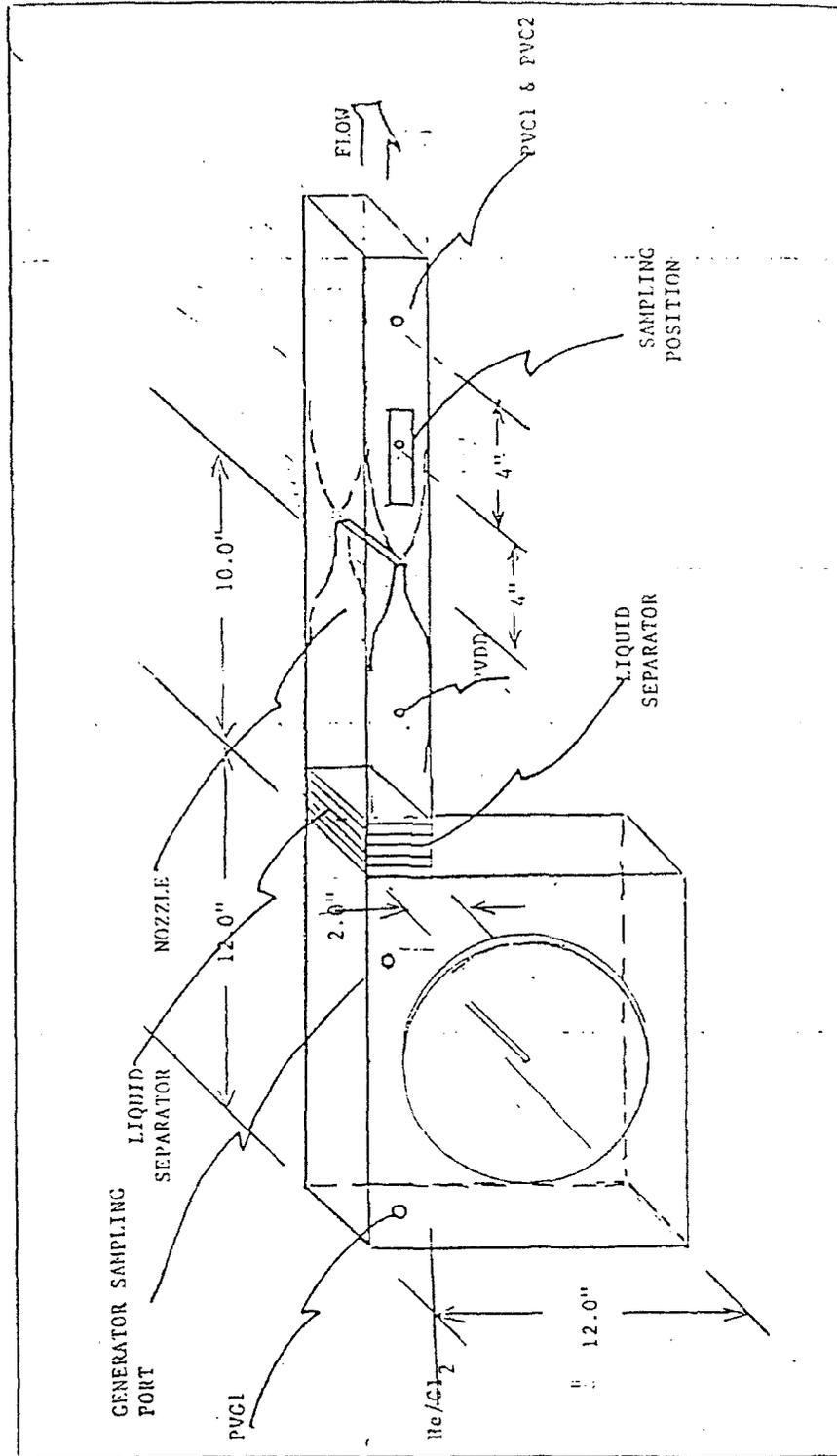


Figure 1. Schematic of "Pee Wee" COIL and Supersonic Flow System

collector particle is $0.968E05$ /s. The size distribution initially monodisperse at size 0.002 microns will broaden with time. The time constant for depletion of the 0.002 micron size is 10.3 microseconds. The net result of coagulation will be a shift in the peak of the size distributor to center on aerosols of larger diameter.

Detection of homogeneously nucleated droplets in the cavity could also be hindered by the possibility of the re-evaporation of the droplets once formed either in the flow system or during sampling. Re-evaporation is not predicted within the laser system where all physical parameters and relative humidity appear to remain constant following the shock[1]. Re-evaporation could represent a serious problem during sampling and dilution of the laser flow stream, where the degree of isothermality required to eliminate the potential for re-evaporation will be demanding. The ammonium hydroxide subsystem of the sampling probe facility was included to address this problem (see experimental section). There are, however factors such as, coagulation, residual aerosol and the possibility that the droplets will form as ice particles, that will help to alleviate the problem of droplet non-detectability.

EXPERIMENTAL

The initial experimental activity in this study was associated with designing, constructing and operating a sample probe facility capable of extracting and maintaining the integrity of a multi-phase sample as it is drawn from a sometimes chemically reactive supersonic flow. The physical parameters of the supersonic nozzle simulation device and the laser fall within the ranges: flow velocity Mach 0.2 - Mach 2.0, total pressure 2 - 50 torr, temperature 150 - 350K. The partial pressure of water lies within the range 2 - 5% of the total system pressure. It is important, therefore, in the case of detecting homogeneously nucleated aerosols, to ensure that there is no temperature increase as the sample flow decelerates through the sampling probe and that the ambient water concentration remains the same. These criteria are met by coupling the sample probe to an evacuable chamber in which reactor species can be diluted and pressurized under isothermal and carefully humidified conditions and temporarily held for examination. Figure (2) is a schematic diagram of this facility. The ammonium hydroxide subsystem is used to convert any homogeneously nucleated aerosols to solution aerosols as they are withdrawn from the flow system. This conversion provides a means of establishing the existence of homogeneously generated aerosols in the event that the aerosols should encounter an environment where they could reevaporate. Once converted to solution aerosol a residue that can be detected will persist after reevaporation. The methodology for heterogeneously nucleated aerosol detection is based on the principle that such aerosols will leave a residue (dry particle) upon reevaporation. Such aerosols

are dried immediately upon extraction from their source and their dry particle properties are characterized. A knowledge of their hydration properties permits the development of growth models.

During this study no effort was made to design and construct a specialized sampling orifice. Instead, a simple tube (orifice i.d. 0.675") was mounted either in the streamlines of the flow (for large diameter aerosols (>0.01 microns) or perpendicular to the flow through an access port mounted on the side of the flow system for smaller aerosols (<<0.01 microns).

The Mobile Aerosol Sampling System (UMR-MASS), figure (3), and the sampling probe facility described above were employed to characterize aerosols in terms of their total number density, size distribution and/or hydration properties. The experimental sequence of events is as follows. (For a more detailed description of the UMR-MASS methodology see references 5-9)[5-9].

(a) Number Density and Size Distributions -- The particulate stream emanating from the sampling probe and dilution facility is sent through, initially, an alcohol counter to examine the total size range of aerosols sampled and determine the total number density. For a size distribution the aerosol is charged with a bipolar charger and is then sized via its electric mobility. The EAC passes only particles contained in a narrow size range (range width is typically 10% of mean size). A specific particle size is set by the EAC voltage. This voltage and hence the selected particle size is variable. The resulting monodisperse particle stream is then passed through the alcohol counter where it is exposed to a fixed supersaturation of alcohol of sufficient magnitude to condense alcohol on the particle, thus forming aerosols large enough to be detected by the optical particle counter, OPC, sub system of the counter. The enlarged aerosols are passed through the OPC where individual aerosols are counted and the aerosol concentration is determined.

(b) Hydration Properties -- For particulate hydration property measurements, specifically the critical supersaturation spectrum (i.e. the minimum supersaturation of water that will cause condensation of water onto a particle to produce a continuously growing droplet presented as a function of particle diameter;) is measured. The particulate stream is first passed through an EAC to select a specific size element of the sample distribution. The resulting monodisperse element is then passed through a saturation chamber holding water vapor at 100% relative humidity. As the aerosols pass through the chamber they deliquesce and increase in diameter. This growth is monitored by passing the droplets, after they exit the saturator, through a second EAC, the output of which can be monitored by the alcohol counter. Should a particulate prove difficult to hydrate (i.e. require significant supersaturations of water a continuous flow diffusion chamber (CFD) is employed.

The setup and operation of the laser device and supersonic nozzle facility

Figure 2. Schematic of the Sample Conditioning system.

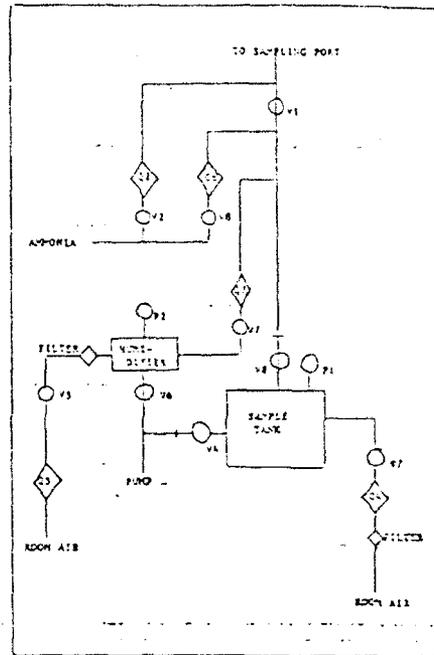
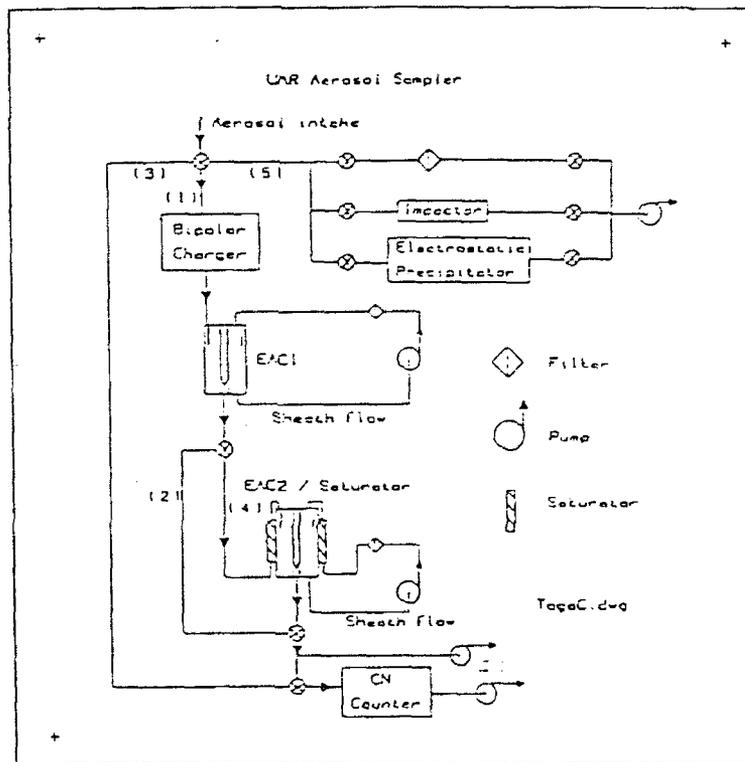


Figure 3. Schematic of UMR-MASS (Mobile Aerosol Sampling System)



is not described in this report. Detailed descriptions of this activity can be found elsewhere [10-12].

RESULTS

Aerosol size distribution data determined with the use of an inversion application [13], show the mean dry diameter of aerosols generated in the oxygen-singlet-delta generator to be approximately $0.03 \mu\text{m}$, see Table 1. Typical size distributions for the generator and cavity and data gleaned from such a distribution can be seen in Figures 4 and 5, respectively. Based on the average dry aerosol diameter of $0.03 \mu\text{m}$ and a 7.0M BHP molarity a nascent wet aerosol diameter was calculated to be $0.047 \mu\text{m}$. Determination of the aerosol particle densities required the adoption of a unit of measure normalizing the results with pressure so that comparisons could be drawn with respect to varying pressure in different sections of the device. The normal particle concentration unit, usually expressed in number of particles per cubic centimeter, was thus modified by dividing the particle concentration by the environmental pressure from which the sample was drawn. This yielded a unit in concentration per torr, allowing comparison of samples without regard for pressure differences. In this manner, the concentration of aerosol particles in a given sample can be determined by finding the product of its new unit of concentration per torr and environmental pressure. Generator aerosol samples typically yielded aerosol densities around 300 particles/cc/torr with excellent reproducibility, see Figure 6. The generator aerosol density does not appear to be related to generator liquid temperature. Cavity aerosol concentrations are typically an order of magnitude lower than the generator aerosol concentrations, ranging from 20 to 90 particles/cc/torr, see Figure 6. Interestingly, although temperature does not appear to play a major role in the variation of the cavity aerosol concentrations, these concentrations seem to increase as the BHP solution is exhausted, see Figure 7.

Further characterization of the aerosol involved determining the critical supersaturation spectra of potential solutes. Theoretical critical supersaturation curves for KOH, KCl, NaCl, and Iodine were generated using the Kohler Equations [14]. Experimental data for each of the salts were in excellent agreement with the theoretical calculated spectra; KCl and NaCl experimental data fell directly on their predicted curves, and KOH was within 10% (instrumental error) of its theoretical curve, see Figures 8 and 9. The theoretical critical supersaturation spectra for Iodine differed from its experimental data, possibly due to approximations in the theoretical model. Nevertheless, both theoretical and experimental spectra for Iodine reside in higher supersaturation regimes than those of the ionic salts as would be expected for substances with low solubilities in water.

TABLE OF AEROSOL SIZE DISTRIBUTION DATA TERMINOLOGY

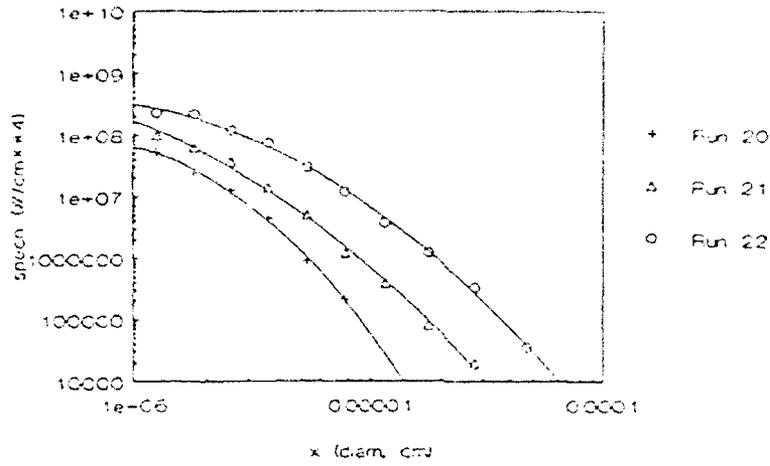
XBAR	Average size (diameter, microns) for this distribution.
XBARA	Diameter corresponding to the aerosol area mean.
XBARV	Diameter corresponding to the aerosol volume mean.
SUMNX	Aerosol's total concentration.
CMASS	Mass of aerosol collected in 60 min at 20 L/min by a cascade impactor which collects particles with diam > 0.2 μ m.
SIZE	Particle diameter for the j-th aerosol family, in cm.
XN	Concentration of the j-th aerosol size family.
SPECN	Aerosol size differential concentration, dC/dD, per unit diameter interval; $\#/cm^{*4}$.
SUM4	Conc. of particles with diam D > 0.04 μ m.
SUM5	Conc. of particles with diam D > 0.05 μ m.
STime	Time in seconds that sample was collected.

TABLE OF SELECT AEROSOL SIZE DISTRIBUTION DATA

Sample	Description	XBAR(cm)	$\#/cc/torr$	STime(s)
Run 11	Gen, no chem	0.541e-5	0.6197	35
Run 12	Gen, 0, mix, 5rpm	0.931e-5	0.2117	40
Run 13	Gen, -18, 45rpm	0.892e-5	60.5363	15
Run 14	Gen, -18, chem	0.290e-5	2926.00	20
Run 16	Cav, -18, chem	0.493e-5	108.459	20
Run 18	Gen, -20, chem	0.345e-5	299.697	2
Run 20	Cav, -10, chem	0.244e-5	18.5714	2
Run 21	Cav, -20, chem	0.274e-5	27.9661	4
Run 22	Cav, 0, chem	0.370e-5	89.1176	4
Run 23	Gen, -10, chem-I2	0.333e-5	307.051	3
Run 24	Gen, 0, chem-I2	0.289e-5	294.600	3
Run 26	Cav, -20, chem-I2	0.303e-5	12.9796	4

Table 1

Aerosol Size Distribution Cavity



Aerosol Size Distribution Generator

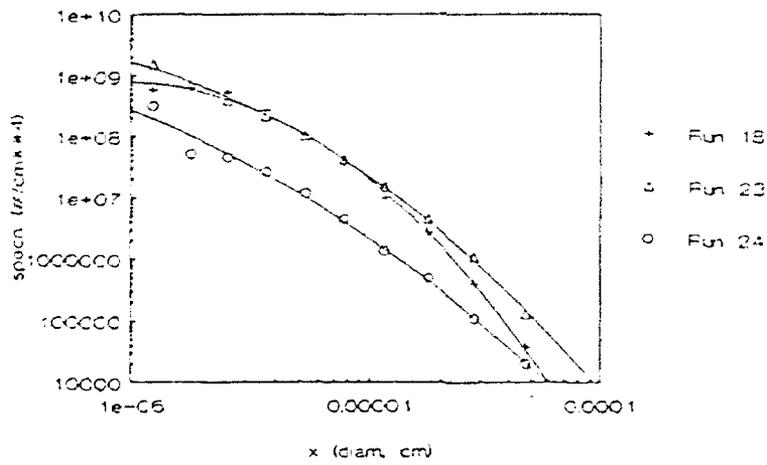
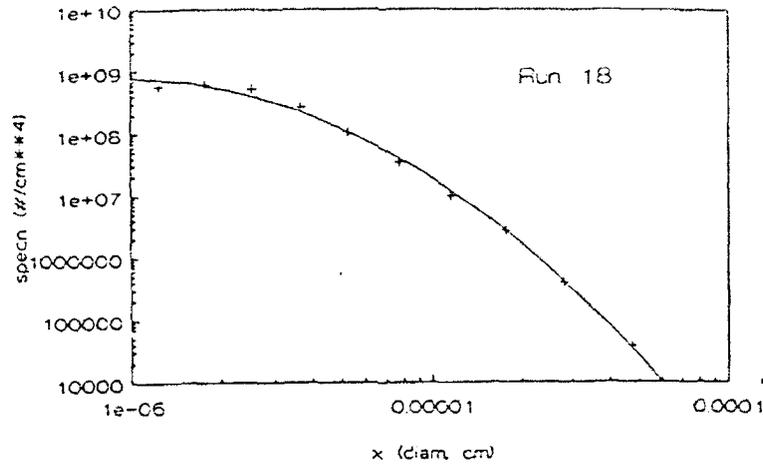


Figure 4

Aerosol Size Distribution Generator



Run 18, decay adjustment (sampling tank):

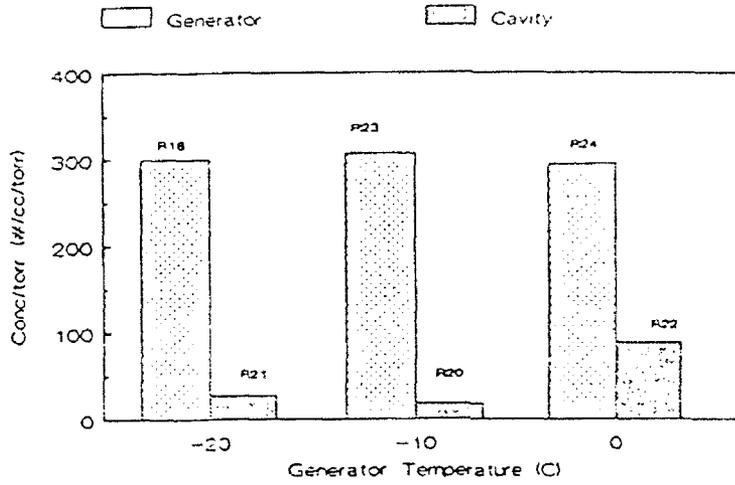
XBAR (cm) = 0.3454D-05 SUMNX (#/cm**3) = 1978.
 XBARA(WRT AREA, CM) = 0.4615D-05 XBARV(WRT VOL, CM) = 0.6392D-

CMASS (MICRO-G) = 0.1292
 SUM4(#/CM**3, > .4 uM) = 401.56
 SUM5(#/CM**3, > .5 uM) = 401.56
 SIGMA = 0.3060D-05
 CCMASS (TOTAL MASS, MICRO-G) = 0.3246

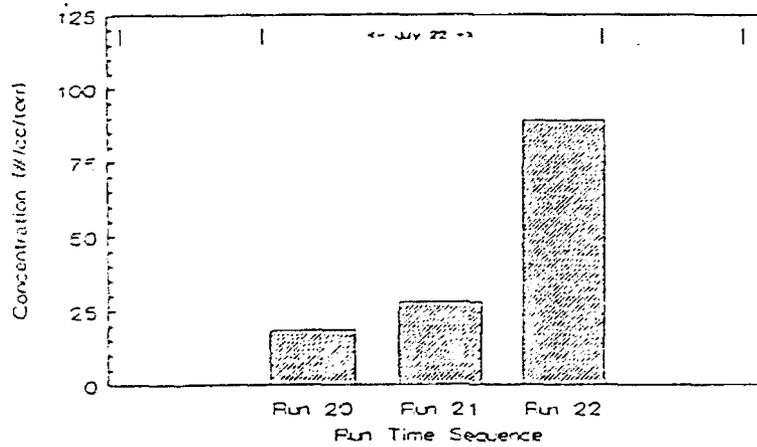
J	POWER	XN (#/CM**3)	SIZE	SPECN	CUMUL
1	1.20	0.310D+03	0.124D-05	0.585D+09	0.216D+10
2	1.20	0.407D+03	0.177D-05	0.630D+09	0.158D+10
3	1.20	0.496D+03	0.254D-05	0.530D+09	0.947D+09
4	2.02	0.363D+03	0.365D-05	0.265D+09	0.417D+09
5	2.64	0.213D+03	0.528D-05	0.104D+09	0.152D+09
6	2.93	0.110D+03	0.774D-05	0.351D+08	0.481D+08
7	3.30	0.494D+02	0.115D-04	0.993D+07	0.131D+08
8	3.74	0.224D+02	0.177D-04	0.270D+07	0.313D+07
9	3.74	0.582D+01	0.282D-04	0.393D+06	0.430D+06
10	3.74	0.711D+00	0.473D-04	0.372D+05	0.372D+05

Figure 5

Aerosol Conc./torr in Gen & Cav. vs. Gen. Temp.

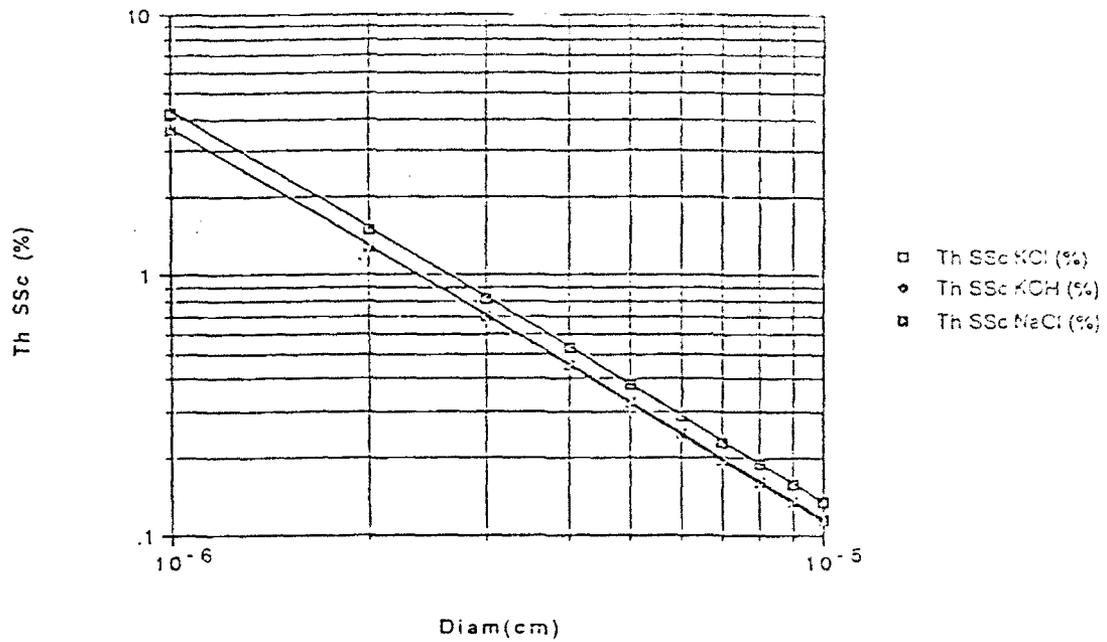


Cavity Conc/Torr vs Run Time Sequence



Figures 6 and 7.

Theoretical SSC of KCl, KOH, and NaCl
vs. Diameter of Particle



Theoretical and Experimental SSC of KOH
vs. Diameter of Particle

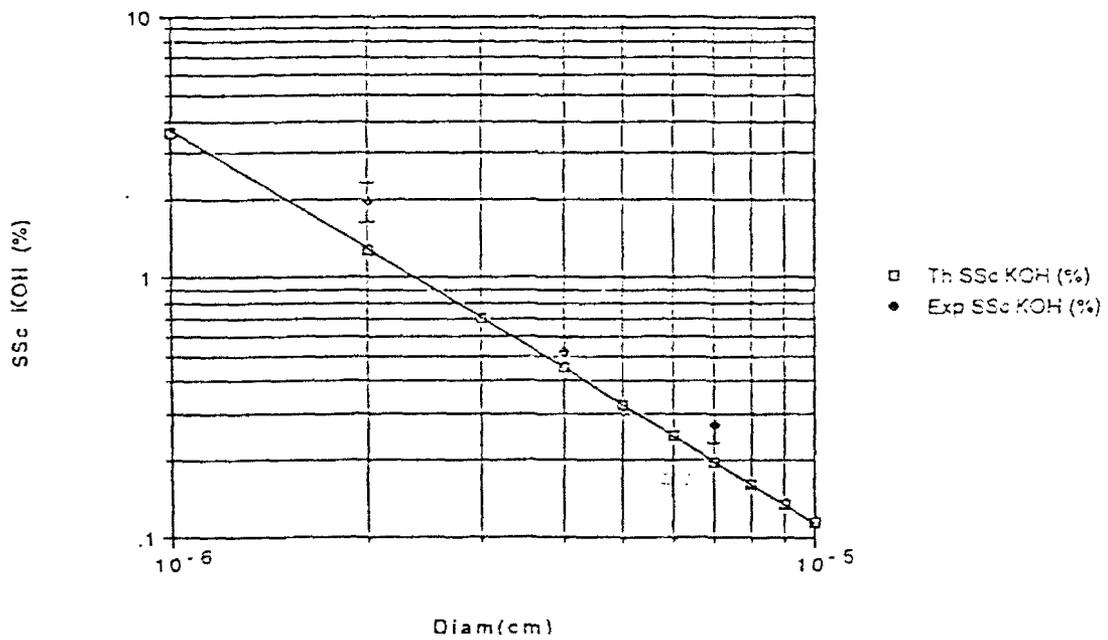
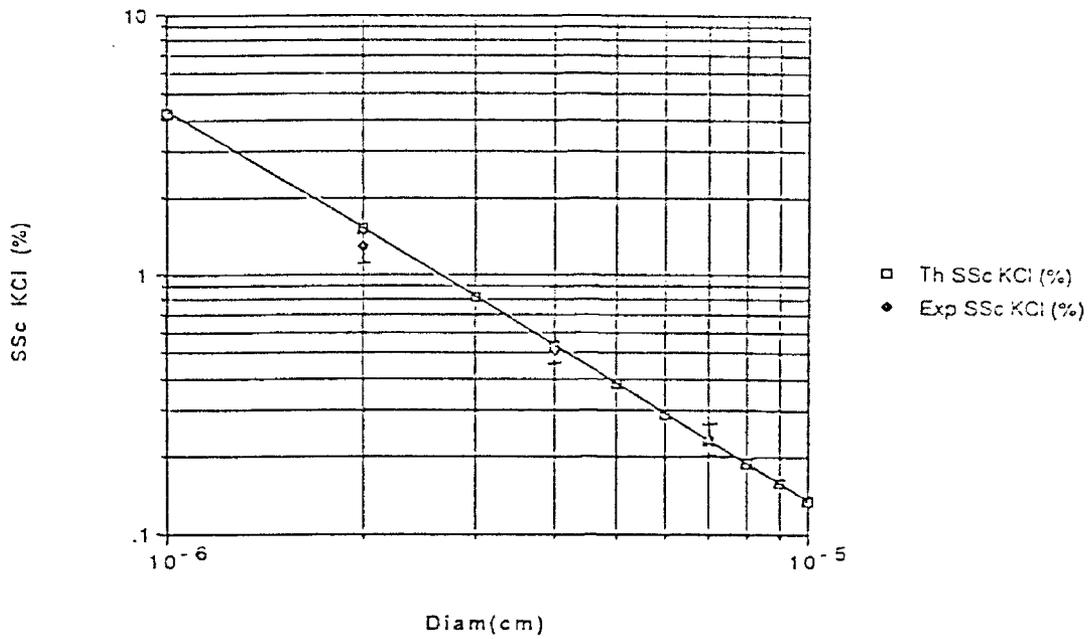


Figure 8.

Theoretical and Experimental SSc of KCl
vs. Diameter of Particle



Theoretical and Experimental SSc of NaCl
vs. Diameter of Particle

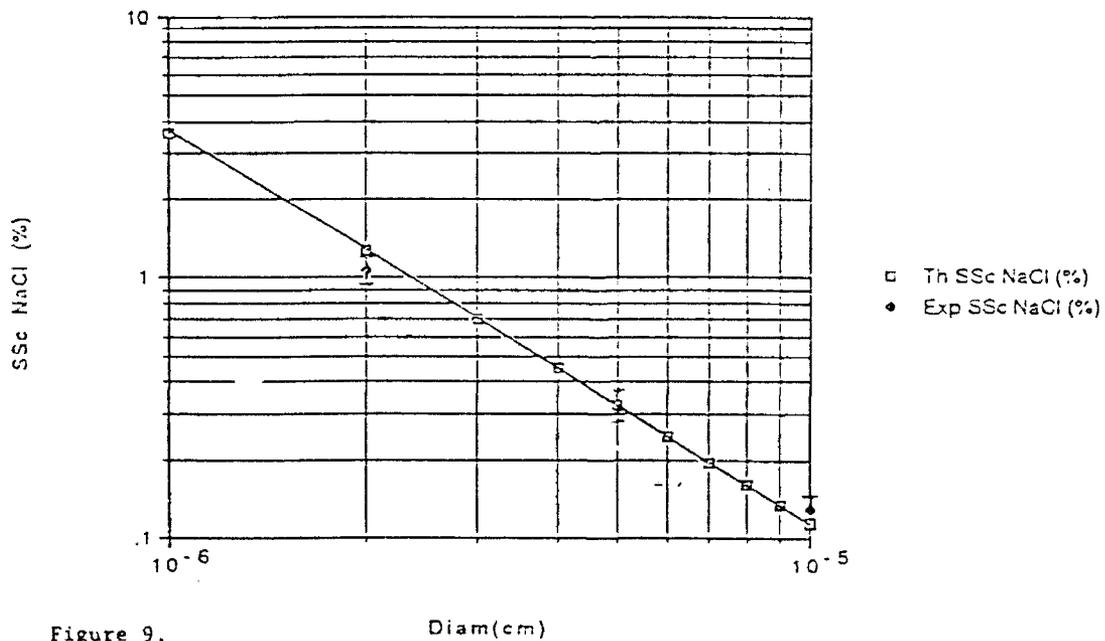


Figure 9.

DISCUSSION

The data acquired with the UMR-MASS was limited by the amount of time the UMR-MASS could be deployed and operated at the Phillips Laboratory based on the funding of this activity through our AFOSR RIP grant. In order to capitalize on this small window of opportunity (10 days) we used this AFOSR sponsored Summer Research Program Faculty Associateship and Graduate Student Associateship to prepare for the arrival of the UMR-MASS, to assist in the data acquisition with the UMR-MASS, to make independent laboratory based measurements (hydration properties) and to interpret the results of this preliminary study.

Homogeneous Nucleation - If the pressure ramp observed; when the three COIL devices were operated with high partial pressures of water vapor; was due to pure homogeneous nucleation down stream of the laser's supersonic nozzle, it should be possible to simulate this event in a supersonic flow system employing the same nozzle, using appropriate mass flows of carrier gas and water vapor. Repeated attempts were made to simulate the pressure ramp in this manner and in all cases no ramp was observed. During the attempts to simulate the pressure ramp cavity gas samples were withdrawn and analyzed with the UMR-MASS for the presence of aerosol. The background aerosol concentration in the supersonic flow system were extremely low < 0.1 particles/cc. and no significant concentrations of aerosol over and above the background count were observed. The logical interpretation of these results are that the pressure increase cannot be a simple gas dynamic effect brought about by homogeneous nucleation.

A re-investigation of the calculations upon which the homogeneous nucleation rational is based is probably warranted at this time. Also, it would be worthwhile to attempt this simulation experiment again but in this case utilizing the Phillips laboratories in-situ water diagnostic to accurately follow the water vapor concentration profile in the supersonic flow facility. A mass spectrometer attached to the flow system would also be a valuable diagnostic to investigate both molecular cluster formation and whether or not charged species emanate from the generator. A justification for such measurements is given later in the discussion.

Heterogeneous Nucleation - This work represents the first experimental data characterizing particulates in the flow stream of a COIL. The typical concentrations measured in this study correlate well with those used in a Mie scattering model to account for non-saturable losses in COILs[15]. As expected the generator is an excellent source of heterogeneously nucleated aerosol. The preliminary data using the UMR-MASS to characterize these aerosols suggests that this is an ideal tool with which to investigate the generation characteristics and fate of these aerosols. These results indicate that heterogeneously nucleated aerosols are mechanically generated during the explosion of gas bubbles on the basic hydrogen peroxide (BHP)/chlorine interface and that certain generator operating parameters such as disc velocity and degree of neutralization

of the BHP can strongly influence the aerosol production rate. To understand the role these aerosols play in the performance and efficiency of a COIL device a test plan should be devised and appropriate measurements with the UMR-MASS performed. This data acquisition effort will require the design and construction of extra sample accumulators, and of, an orifice with high sampling efficiency in the supersonic cavity region. Additional sampling ports will also be required before and after the liquid separator plates and cold traps (if applicable) and beyond the cavity region close to the pumps. Data should also be acquired on an independent device such as RADICL. The opportunity for intercomparison of data from two separate devices would be of great value. The impact of these aerosols on iodine dissociation and excitation should also be explored. The hydration property measurements indicate that iodine coated aerosol will have higher critical supersaturations than those of the simple ionic salt aerosols.

These preliminary experiments answer many questions about aerosols in COILs and have led to recommendations for further study particularly in the case of characterizing the role of heterogeneously nucleated aerosols. The pressure ramp and potential condensation shock issue, however, have not been resolved. It is clear from the data on number density, size distribution and hydration properties of heterogeneously nucleated aerosols that these species, as characterized in these preliminary experiments, alone could not condense enough water within their cavity residence time to release the required amount of heat. There is a possibility, however, that the true number of heterogeneously nucleated aerosols emanating from the generator has been underestimated in the experiments thus far. Should the true number of aerosols exceed the measured number by several orders of magnitude then it would be possible to account for the pressure ramp via an enthalpy of condensation rational. The UMR-MASS, although state of the art in detectability, has a lower limit in aerosol diameter detection of 0.01 microns. This limit is governed by the design of the aerosol counter. A close inspection of the size distribution plots, figures (4) and (5), do not show that the size distribution is single mode or that the peak in the distribution has been clearly met. There is no a priori reason to expect a single mode distribution here. For combustion aerosols from jet engines for example, a single mode distribution is always found[7], but in comparison a typical atmospheric aerosol distribution is far more complex and tends to have a significant number density beyond the limit of detectability of CNC counting systems[14]. The mechanical generation of aerosols via bubble explosions on the liquid/gas interface has been examined previously[15] and may lend credence to the underestimation of aerosol concentrations. The size of the aerosols produced is directly related to the initial bubble diameter. Submicron dry particles in the size ranges of interest in this problem have been detected from sea water foam. Serious thought, therefore, should be given to improving the technique to detect and count smaller

diameter aerosols.

One other potential mechanism to explain the pressure ramp through condensation on a high concentration of nucleating species ($>10^8/\text{cc}$) is to propose the presence of charged species in the generator effluent. One method for examining this possibility would be to introduce an electric field perpendicular to the flow stream in the subsonic flow region of the duct.

The valuable original research data and inevitable questions arising from this preliminary study open the door to further discussion on the approach for future experiments some of which could be clearly identified as tasks for a 1992 RIP proposal.

CONCLUSIONS

Heterogeneously nucleated aerosols are present in COILs. They are mechanically produced in the generator and their production rate and size are governed by generator operating parameters. The relationship between these parameters and aerosol production should be explored in detail. The limits of detectability of aerosols using the UMR-MASS in its current state of the art configuration may have led to an underestimate of the total concentration of aerosols produced in the generator. As a result of the studies to date, design up grades should be made to the sampling probe system and counting system to customize the UMR-MASS to this COIL investigation.

The inability to simulate the pressure ramp in the supersonic nozzle flow system indicates that the pressure ramp is not a simple gas dynamic effect arising from the enthalpy of condensation associated with a homogeneous nucleation event. The calculations upon which the homogeneous nucleation rational are based should be re-examined.

The principles of operation and performance of the spray generator concepts should be examined to determine whether submicron aerosol formation will occur in these devices and whether it will affect their predicted performance.

7.0 References.

- (1) Private Communication, D. Plummer RDA, 19 July 1991.
- (2) RADICL Data presented at a briefing on the condensation shock phenomena by D. Plummer RDA, March 1992.
- (3) Phillips Laboratory Small Scale COIL data presented by C. Helms, July 1992.
- (4) Fundamentals of Gas Dynamics, Editor H.W. Emmons, Princeton University Press, 1958.
- (5) Alofs, D.J., "Performance of a dual-range cloud nucleus counter", J. Appl. Meteor. 17, 1286-1297 (1978).

(6) Alofs, D.J., M.B. Trueblood, D.R. White, and V.L. Behr, "Nucleation experiments with monodisperse NaCl aerosols", J. Appl. Meteor. 18, 1106-1117 (1979).

(7) The influence of jet fuel sulfur impurities on the hydration properties of jet fuel-generated combustion aerosols, P.D. Whitefield, M. Trueblood and D.E. Hagen, presented at the A.A.A.R. 10th annual meeting October 1991.

(8) Alofs, D.J. and M.B. Trueblood, "UMR Dual Mode CCN Counter (Modes: CFD plus Haze)," J. Rech. Atmos. 15, n' 3-4, 219-223 (1981)

(9) Hagen, D.E., M.B. Trueblood, and J. Podzimek, "Combustion Aerosol Scavenging", Atmos. Env. 25A, 2581 (1991).

(10) Benard, D.J., McDermott, W.E., Pchelkin, N.R., and Bousek, R.R., "Efficient Operation of a 100 Watt Transverse Flow Oxygen-Iodine Chemical Laser", Appl. Phys. Letters 34, 40-41, (1979).

(11) Richardson, R.J., Wiswall, C.E., Carr, P.A.G., Hovis, F.E., and Lilienfeld, H.V., "An Efficient Singlet Oxygen Generator For Chemically Pumped Iodine Lasers", J. Appl. Phys. 52, 4962-4969, (1981).

(12) Oxygen Iodine Supersonic Technology Program, Part III - Management and Technical Proposal, MDRL/MDAC Report No. D9299-953003, June 1982.

(13) Hagen, D.E., and D.J. Alofs, "A Linear Inversion Method to Obtain Aerosol Size Distributions from Measurements with a Differential Mobility Analyzer", Aerosol Sci. and Tech. 2, 465-475 (1983).

(14) Whitefield P.D., D.E. Hagen and M.B. Trueblood, "A field Sampling of Jet Exhaust Aerosols." Particulate Science and Technology, (submitted August 1992)

(15) H.R. Pruppacher and J.D. Klett Microphysics of Clouds and Precipitation, D.Reidel Publishing Co., Dordrecht, Holland 1978

(16) RDA Memo 87-A/K-3-02-752 Particulates/Aerosols in Oxygen/iodine Lasers, P.G. Crowell, January 1989.

**INSTALLATION OF OSCILLATING NEURAL NETWORK ALGORITHMS
INTO A KHOROS TOOLBOX**

**Carla C. S. Williams
Assistant Professor
Department of Physics and Engineering**

**Fort Lewis College
Durango, CO 81301**

**Final Report for:
Summer Research Program
Phillips Laboratory**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.**

August 1992

INSTALLATION OF OSCILLATING NEURAL NETWORK ALGORITHMS INTO A KHOROS TOOLBOX

Carla C. S. Williams
Assistant Professor
Department of Physics and Engineering
Fort Lewis College

Abstract

Many of the Air Force laboratories are interested in providing an environment supporting multidimensional signal processing research using neural network algorithms for pattern recognition and classification. This paper presents the results of an on-going project to develop, implement, and test a number of generic neural network algorithms that could be used by an application engineer with limited domain expertise in neural computation. The current focus of the project is on the simulation of coupled oscillatory neural networks which hold promise in exploiting the massive parallelism offered by newer hardware technologies. Several routines were implemented under the Khoros software development system which provides tools for integrating new programs to be used within a visual programming environment which also includes a large number of predefined signal and image processing routines.

INSTALLATION OF OSCILLATING NEURAL NETWORK ALGORITHMS INTO A KHOROS TOOLBOX

Carla C. S. Williams

BACKGROUND

Many of the Air Force laboratories are currently investigating applications that have the potential to benefit from the advantages offered by neural networks in pattern classification. Last summer [1], several neural algorithms that had previously been developed at the Phillips Laboratory [2] were implemented under Khoros, a generalized visual programming environment developed at the University of New Mexico. The goal of the project this summer was to examine several algorithms for simulating coupled oscillatory neural networks and to implement them in Khoros Toolbox. The Khoros software system was chosen because it provides an intuitive graphical user interface to routines while allowing the user access to a large assortment of predefined signal and image processing routines.

Coupled Oscillatory Neural Networks

Neural networks presently offer a way to exploit the large scale parallelism offered by new hardware technologies in complex situations where knowledge cannot be encoded directly and where speed is important. However a more general-purpose neural computation machine is needed which will model causality permitting mixed modes of operation using both formal logic and learning from examples. One approach is to encode non-equilibrium "state" as oscillations in a set of coupled non-linear elements similar to those found in the cerebral cortex [3]. Examples of such systems have been tested and it is generally believed that complex temporal relationships can be encoded and learned using such systems of oscillators. For example, olfaction has been simulated as a process for identification and classification. This type of behavior can be exploited to model causality which may be particularly useful for detecting faults and controlling complex machinery such as satellites and autonomous vehicles.

Khoros

The Khoros system [4], implemented at the University of New Mexico as an ongoing research project, is an integrated software development environment for multidimensional information processing and data visualization. The system consists of a number of specialized tools for building programming applications. These include an editor for specifying the user interface for applications, code generators to create all the user interface code for both graphical and command-line user interfaces, and source configuration and management tools. Applications created with the tools and supported as part of the Khoros system include over 290 programs for processing image/signal information and for data display and manipulation. The most significant component of the Khoros system is the visual language, *cantata*, which provides a graphical programming interface to the applications.

Visual languages, which combine the disciplines of interactive graphics, computer languages, and software engineering, have the potential to dramatically enhance the communication between human and computer by providing a more intuitive interface to programming. As used in *cantata*, programs are constructed as *data flow graphs*, or block diagrams, which are already familiar to scientists and engineers. The nodes represent block processes of large granularity, such as filtering or classification algorithms, so that solutions to problems can be configured at a high level without the user needing intimate knowledge of the details of implementing a specific algorithm.

Figure 1 presents a *cantata* program workspace using two routines that were developed for this project, the Baird oscillating training algorithm (*osc_t*) and classifier (*osc_c*). Four input matrices are fed into the *osc_t* module to train the network. The resulting output (a weight matrix) provides input to the *osc_c* module along with the vector(s) to be classified. Classification statistics are displayed in an ASCII text window produced by the *xvviewer* module (see Figure 2a). The *xprism3* module is used to view a three-dimensional plot of the oscillations (see Figure 2b).

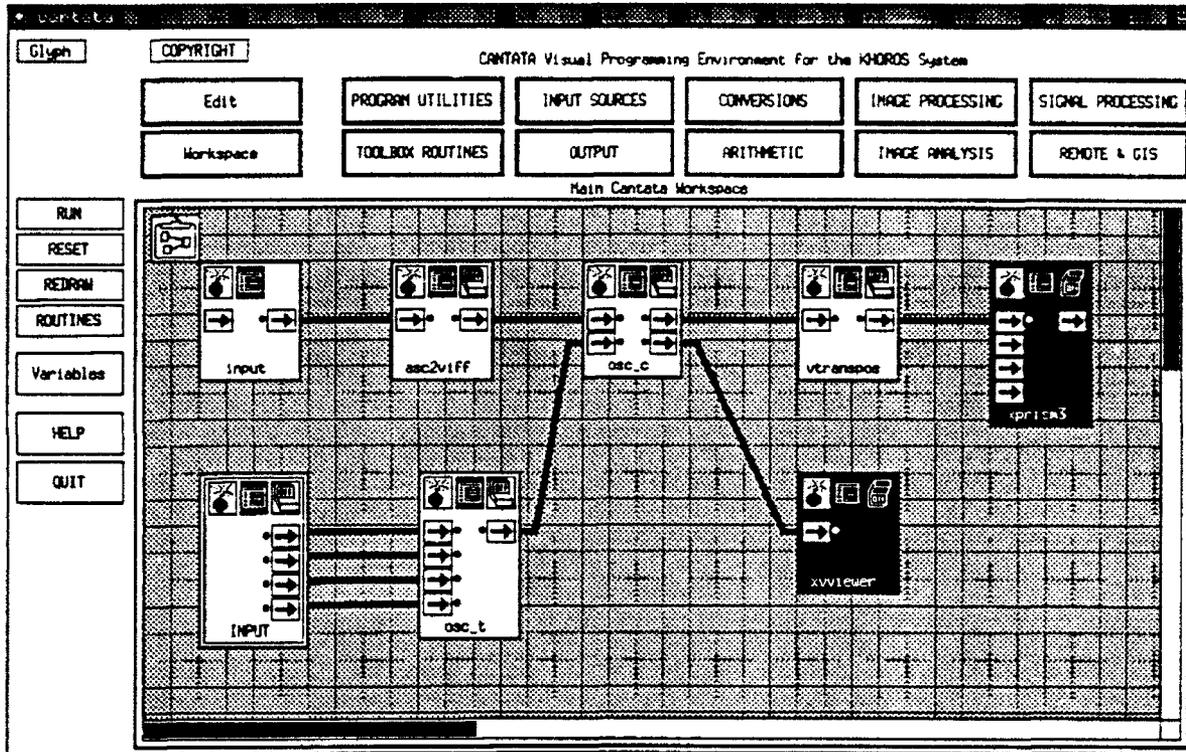


Figure 1 Cantata workspace using Baird Oscillating Neural Network.

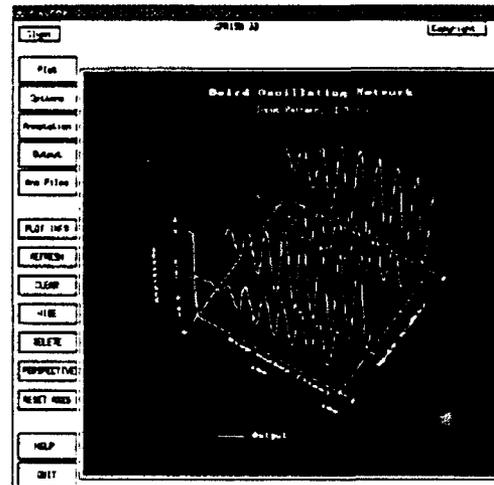
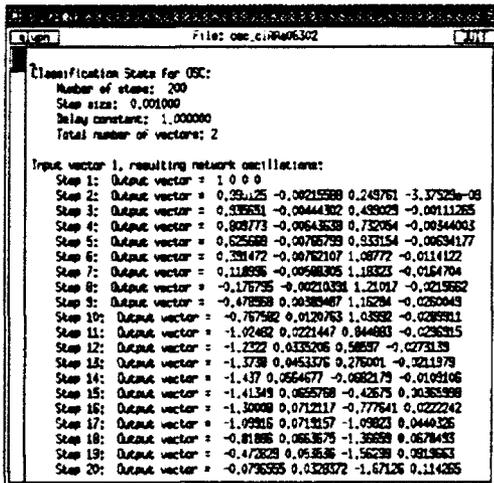


Figure 2a Classification statistics. Figure 2b Plot of classification output.

Output of Baird Oscillating Neural Network - Classification.

APPROACH

To create a Khoros Toolbox the following steps were identified (which are described in more detail below):

- (1) Develop code for the routine,
- (2) Create and test the routine in Cantata,
- (3) Install the routine in a Toolbox.

Develop Code for the Routine

The C code program which simulates the Baird oscillating neural network for this project was developed by John Greenfield and is described elsewhere [5].

Once the program code that contains the functionality for the routine has been developed and tested, it must be structured for inclusion into Khoros as a *vroutine*. Khoros *voutines* for information processing are typically written as a main program which then calls a library routine. The structural organization I chose for the neural routines is as follows:

MAIN: Input/output of data and parameters plus error checking.

LIBRARY: Preprocessing of data and memory allocation. Training/testing of input data vectors and output of intermediate results.

To be compatible with the image/signal processing routines currently available within Khoros, new modules need to use the same data structure for input/output data. The *Khoros Visualization/Image File Format (VIFF)* and internal data structure was designed to facilitate the exchange of data, not only between Khoros routines, but also between researchers exchanging data. Data input/output for a module is done in the main driver program (written in C) using the Khoros utilities to read, write, and manipulate the VIFF data structure. The VIFF structure is organized as 1024 bytes of header followed by map(s), location, and image data. An image is a two-dimensional, multiband array of data values.

Input training and testing data for oscillating neural network routines consist of a set of matrices and the input vectors. To store this information in the VIFF structure, I chose to represent the matrices and vectors as multiband images which were input by the main program. The library routine would then extract the data from the image structure in the correct format for processing.

Create and Test the Routine in Cantata

The steps in creating a Khoros *vroutine* include:

- (a) create the graphical User Interface Specification (*.pane) file,
- (b) complete the Program Specification (*.prog) file,
- (c) generate code, documentation, and Make files.

The graphical user interface (GUI) utilized by processing routines implemented under Khoros consists of a *subform/pane*. Each subform contains the information for a group of related routines from the Khoros libraries. Figure 3 shows the 'Oscillating Neural Networks' subform created for this project. The pane for the 'Baird Oscillating Neural Network - Classification' program, *osc_c*, is shown on the right side of the subform. The GUI is defined by a *User Interface Specification (UIS)* which is an ASCII file composed of a sequence of lines following a strict syntactical structure, each of which describes a separate part of the user interface. The parameters in each line control the data type, position, size, default values, variable names, command-line switches, and action controls for that item. Figure 4 gives an excerpt of the UIS for the 'Oscillating Neural Networks' subform with the *osc_c* pane. All the input parameters needed to configure the neural network for classification are listed, their values to be specified by the programmer in when the module is selected for use.

A UIS may be created in one of two ways, or (as often happens) a combination of both. The Khoros graphical UIS editor *composer* can be used by an application developer to interactively select appropriate options describing the desired user interface, which are automatically translated into the UIS

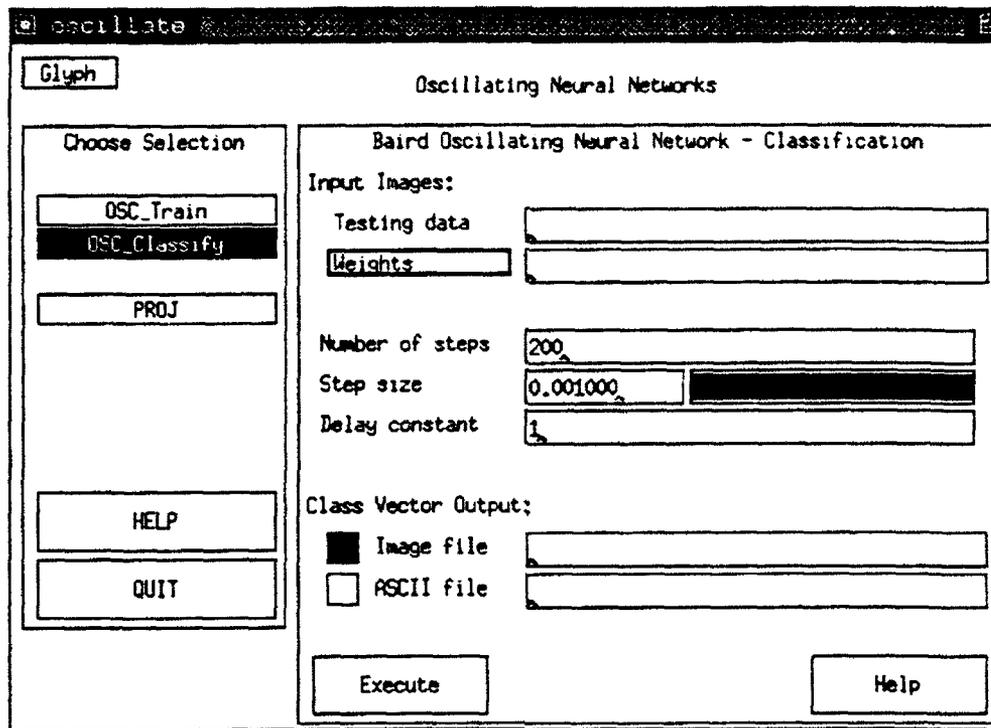


Figure 3 Subform for Oscillating Neural Network algorithms.

file format, then interpreted and displayed on the screen. Alternately a standard text editor can be used to create the UIS file directly (often starting with a previously defined UIS from a similar application). In this case, a prototype of the graphical interface can be interpreted and then displayed by the UIS preview program and modified until its appearance is as desired. Syntax checking is automatically done on each entry for the new form and erroneous or illegal specifications are rejected.

The C code for the main driver and library are included in a *Program Specification* file, which may be created with the assistance of **composer** or by using a standard text editor. In this file, the application developer provides additional information that cannot be included in the UIS, such as program documentation, syntax for library calls, as well as the specific code segments that are to appear in the application program. Figure 5 gives the Program Specification for `osc_c`. This is used, along with the UIS (Figure 4), by the code generators to create the complete code for the new module.

```

-F 4.2 1 0 170x7+10+20 +35+1 'CANTATA Visual Programming Environment for the
KHOROS System' cantata
-M 1 0 100x40+10+20 +29+1 'Oscillating Neural Networks' oscillate
-G 1 20x38+1+2 +2+0 'Choose Selection'
-g 1 1 18x1+1+2 'OSC Train'
-g 1 0 18x1+1+3 'OSC Classify'
-g 1 0 18x1+1+5 'PROJ'
-H 1 18x2+1+11 'HELP' 'guide help' $OSCILLATE/doc/cantata/subforms/help/Overview
-Q 1 0 18x2+1+13 'QUIT'
-E
...

-P 1 0 80x38+22+2 +2+0 'Baird Oscillating Neural Network - Classification' osc_c
-b +0+1 'Input Images:'
-I 1 0 0 1 0 0 50x1+2+2 +0+0 ' ' 'Testing data' 'testing data' il
-I 1 0 0 1 0 0 50x1+2+3 +0+0 ' ' 'Weights' 'input weights' wt
-i 1 0 0 1 0 50x1+1+5 +0+0 2 2 1000 'Number of steps' 'number of steps' maxos
-f 1 0 0 1 0 50x1+1+6 +0+0 0 1 0.001 'Step size' 'step size' h
-f 1 0 0 1 0 50x1+1+7 +0+0 2 2 1 'Delay constant' 'time delay constant' tau
-b +0+9 'Class Vector Output:'
-O 1 0 1 1 0 0 50x1+2+10 +0+0 ' ' 'Image file' 'class output vectors' o
-O 1 0 1 0 0 0 50x1+2+11 +0+0 ' ' 'ASCII file' 'classification stats' f
-R 1 0 1 13x2+1+13 'Execute' 'do operation' osc_c
-H 1 13x2+39+13 'Help' 'man page for osc_c' $OSCILLATE/doc/manpages/osc_c.1
-E
...

-E
-E

```

Figure 4 UIS file for Oscillating Neural Networks subform / osc_c pane.

The code generators for Khoros act as programmer's assistants by providing aid in the tasks that prove to be invariant across applications. These include generation of modular code to extract user input from the user interface, and creation of the C structures needed to mediate between the application program and the user interface, plus the code to initialize these structures. Also generated are the generic application drivers, as well as consistently formatted manual pages and usage statements.

Continuing with the Baird Oscillating Neural Network - Classification algorithm as an example, the code generator ghostwriter takes as input the UIS file (osc_c.pane) and Program Specification (osc_c.prog) and produces the following files:

```

-AUTHORS
J. Greenfield and G. Williams
-AUTHORS_END

-SHORT_PROG_DESCRIPTION
Baired Oscillating neural network classification algorithm.
-SHORT_PROG_DESCRIPTION_END

-SHORT_LIB_DESCRIPTION
Baired Oscillating neural network classification algorithm.
-SHORT_LIB_DESCRIPTION_END

-NAME_LONG_DESCRIPTION
I osc_c
is the classification portion of a pair of programs that simulate an
Oscillating neural network, using the Normal form projection algorithm
from [1]. The corresponding training program (osc_t) produces both the
excitation and inhibition weight matrices required by this classification
algorithm. This program retrieves the memorized oscillatory pattern that
is the smallest Hamming distance from the initial input pattern. Thus a
single static input pattern produces a stable dynamic oscillation.
.SH
Reference:
LP
[1] S. Baired and F. Eeckman, "A Normal Form Projection Algorithm for
Associative Memory."
.SH
Program Parameters:
LP
TESTING DATA (t): a VIFF image that contains the input vectors to be
classified. Each vector is defined as an image band where the total number
of elements in each vector (band) is a number of rows * number of columns.
LP
WEIGHTS (wt): a VIFF image that contains the weights computed by the
training program (osc_t). This is used to classify the testing data.
LP
NUMBER OF STEPS (maxst): the iteration limit used to stop
oscillation of the network (default is 1000).
LP
STEP SIZE (h): step size for the differential equation solver.
LP
DELAY CONSTANT (tau): time delay constant.
LP
IMAGE FILE (e): a VIFF image which is used to store the output vectors
computed by the oscillating network. Each vector is defined as a band with
the vector elements for each oscillation step contained in a single row. The
order and number of output vectors corresponds to that of the input vectors.
LP
ASCII FILE (f): the filename of an ASCII file to which the classification
stats will be printed.
During classification, the program will print the output of the network
while it is oscillating.
If the ASCII FILE is not specified, the classification stats will be
printed to standard output.
-NAME_LONG_DESCRIPTION_END

-NAME_EXAMPLES
osc_c -i input.vv -maxst 100 -h 0.001 -tau 1 -wt weight.vv -e classout.vv
-NAME_EXAMPLES_END

-NAME_RESTRICTIONS
All input images MUST be of data storage type FLOAT.
The output "vector" image is data storage type FLOAT.
-NAME_RESTRICTIONS_END

-NAME_SEE_ALSO
osc_t(1)
-NAME_SEE_ALSO_END

-NAME_LONG_DESCRIPTION
I osc_c
is the library classification portion of a pair of programs that simulate
an Oscillating neural network, using the Normal form projection algorithm
from [1]. The corresponding training program (osc_t) produces both the
excitation and inhibition weight matrices required by this classification
algorithm. This program retrieves the memorized oscillatory pattern that
is the smallest Hamming distance from the initial input pattern. Thus a
single static input pattern produces a stable dynamic oscillation.
.SH
Reference:
LP
[1] S. Baired and F. Eeckman, "A Normal Form Projection Algorithm for
Associative Memory."
-NAME_LONG_DESCRIPTION_END

-NAME_RESTRICTIONS
All input images MUST be of data storage type FLOAT.
The output "vector" image is data storage type FLOAT.
-NAME_RESTRICTIONS_END

-NAME_SEE_ALSO
osc_t(1)
-NAME_SEE_ALSO_END

-USAGE_ADDITIONS
-USAGE_ADDITIONS_END

-INCLUDE_INCLUDES
-INCLUDE_INCLUDES_END

-INCLUDE_ADDITIONS
-INCLUDE_ADDITIONS_END

-INCLUDE_MACROS
-INCLUDE_MACROS_END

-MAIN_VARIABLE_LIST
struct vimage {
    "data," /* input testing data */
    "wtimg," /* input weights matrix */
    "outvec," /* computed output vectors */
    "readimage()",
    "createimage()",
};
int nvec, /* number of testing vectors */
    N, /* number of nodes */
FILE "printdev"; /* print device for ascii output of stats */
-MAIN_VARIABLE_LIST_END

-MAIN_BEFORE_LIB_CALL
if (check_args())

```

Figure 5 Program Specification for osc_c.

```

the output vector oscillations computed for the input testing data.
-LIBRARY_OUTPUT_DND

-LIBRARY_DEF
inc losc_c (data, wimg, outvec, maxoe, h, tau, printdev);

struct wimage {data, *wimg, *outvec;
int maxoe;
float h, tau;
FILE *printdev;
-LIBRARY_DEF_DND

-LIBRARY_CODE
{
int notv, M, sizeT, tv, i, j, k, l, c, index1, index2;
float sum;
float *bvec, *W, *T, *T2, *Ovect;
float *b, *maxoe, *tau;
char *program = "losc_c";

/* Extract testing data from image structure */
bvec = (float *) data->imgedata;
notv = data->num_data_bands; /* number testing vectors */

/* Extract weight matrix from image structure */
W = (float *) wimg->imgedata;
M = wimg->row_size; /* size testing vectors */
sizeT = M * M;
T = W;
T2 = A * W(sizeT);

/* Allocate memory for computed Output vector oscillations */
if ((Ovect = (float *) malloc(M*maxoe*notv*sizeof(float))) == NULL)
{
(void)fprintf(stderr, "%s: Unable to allocate Ovect matrix.\n", program);
return(0);
}
for (i=0; i<(M*maxoe*notv); i++)
Ovect[i] = 0.0;

/* Allocate memory for output vector */
if ((x = (float *) malloc(M*sizeof(float))) == NULL)
{
(void)fprintf(stderr, "%s: Unable to allocate x vector.\n", program);
return(0);
}

/* Allocate memory for intermediate output vector */
if ((nu = (float *) malloc(M*sizeof(float))) == NULL)
{
(void)fprintf(stderr, "%s: Unable to allocate nu vector.\n", program);
return(0);
}

(void)fprintf(stderr, "\nClassification Rate for OSC:\n");
(void)fprintf(stderr, "Number of steps: %d\n", maxoe);
(void)fprintf(stderr, "Step size: %f\n", h);
(void)fprintf(stderr, "Delay constant: %f\n", tau);
(void)fprintf(stderr, "Total number of vectors: %d\n", notv);

/* Process input testing vectors, one at a time */
for (tv=0; tv<notv; tv++)
{
/* Extract individual vector */
b = &bvec[M*tv];
xvec = &Ovect[M*maxoe*tv];

/* Oscillate on an individual vector */
(void)fprintf(stderr, "Input vector %d, resulting network oscillations:
tv=);

.....
/* This code implements the classification part of the first
/* coupled oscillator in Baird's paper
.....

for (i=0; i<M; i++)
x[i] = 0.0;
for (i=0; i<maxoe; i++)
for (l=0; l<M; l++) {
sum = 0.0;
for (j=0; j<M; j++) {
index1 = l*M + j;
sum += T[index1] * x[j];
for (k=0; k<M; k++)
for (l=0; l<M; l++) {
index2 = l*M*M + j*M + k*M + l;
sum += T2[index2] * x[j] * x[k] * x[l];
}
}
nu[i] = h * (tau * x[i] + sum);
if (i==0)
nu[i] = b[i];
}
for (i=0; i<M; i++) {
x[i] = nu[i];
xvec[M*i] = x[i];
}

/* Print iteration stats */
(void)fprintf(stderr, " Step %d: Output vector = ", i+1);
for (i=0; i<M; i++)
(void)fprintf(stderr, "%g", x[i]);
(void)fprintf(stderr, "\n");
}
}
outvec->imgedata = (char *) Ovect;

/* Cleanup */
if (x != NULL) free((char *) x);
if (nu != NULL) free((char *) nu);

return(1);
}
-LIBRARY_CODE_END
-LIBRARY_CODE_DND
-LIBRARY_CODE_END

```

Figure 5 (cont.) Program Specification for osc_c.

Main program source: osc_c.c

Library routine source: losc_c.c

Header file: osc_c.h

Main program manpage: osc_c.1

Library routine manpage: losc_c.3

Template configuration file: osc_c.conf

plus Make utility files: Imakefile and Makefile for compiling the programs.

Install the Routine in a Toolbox

Routines developed to extend the Khoros system should be installed in and maintained in a *toolbox*. A toolbox consists of a set of files that make up the toolbox source tree and specify the operation of the Make utilities. The toolbox directory structure mimics the Khoros home directory, but it may be created in any location. The following steps are needed to use a toolbox:

- (a) create and configure the toolbox,
- (b) install routine into the toolbox,
- (c) set up access for the new toolbox.

Toolboxes are managed through a **Toolboxes** file, which describes all the Khoros toolboxes to be accessed by an individual user. The file must first be created (by copying a provided template) in the user's directory. The **Toolboxes** file is made up of one entry for each toolbox, each entry on a separate line. The **kraftsman** program is used to create the description of a new toolbox, build the toolbox directory structure, and to configure it correctly.

All the files created for a new routine must now be moved into the correct location within the toolbox directory. This is accomplished by the Khoros utility **kinstall**, which uses the locations provided by the programmer in the ***.conf** file. In addition, the toolbox *subform* UIS file (which provides the GUI menu within **cantata** to all the routine *panes* in the toolbox) is created and maintained under programmer supervision. Because of limitations in the current implementation of **kinstall**, the first routine to be installed into a new toolbox requires an extra step to manually move the library code into the **Lib** directory and create the **Imakefiles** and **Makefiles** for the source tree. Figure 6 shows the directory structure created for the Oscillating Neural Networks toolbox **oscillate**.

Finally, the user's environment must be amended to allow access to the new toolbox. This includes adding a path to the executables for the new routines as well as modifying the **cantata** UIS form file which references the new toolbox.

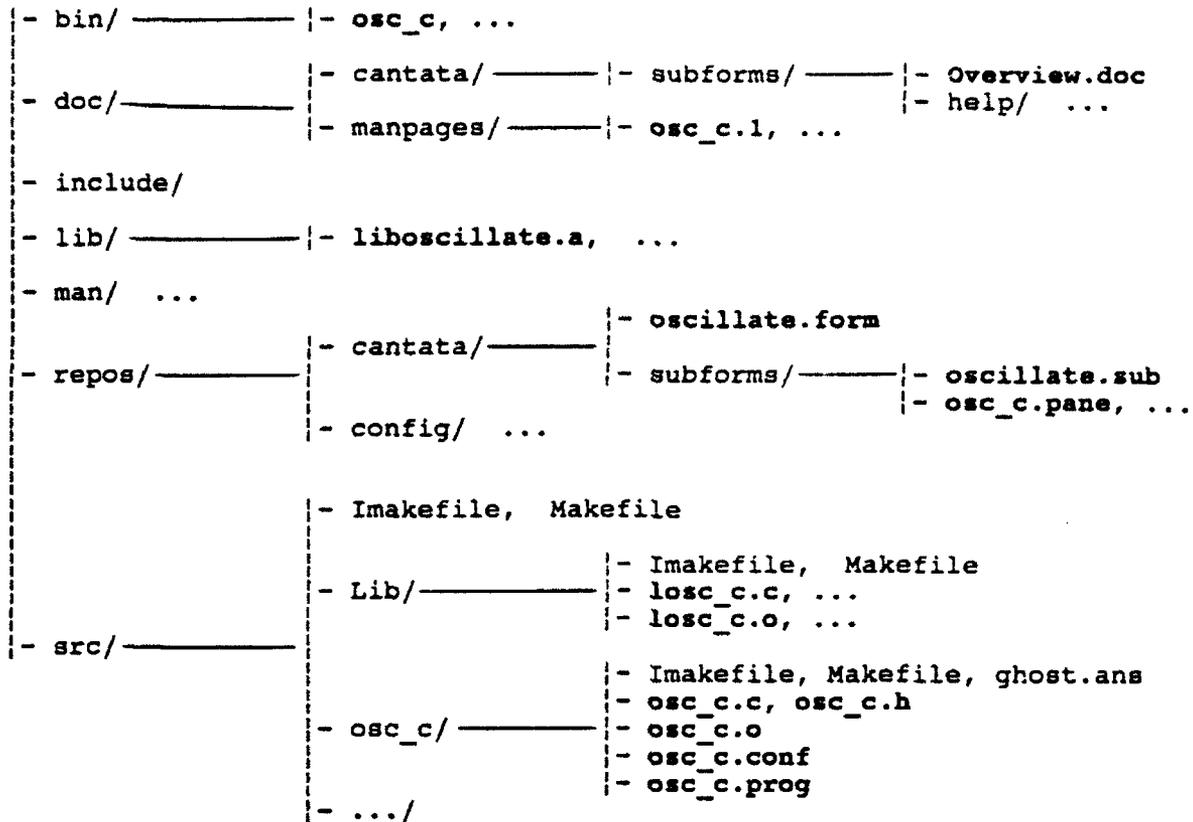


Figure 7 Directory structure for oscillate toolbox.

CONCLUSION

The Khoros system offers many advantages for supporting the desired neural network environment. The cantata visual programming language provides an intuitive design platform conducive to "what if" experimentation and learning. The software development tools facilitate the importation of new algorithms as operator modules within the visual language. The new coupled oscillating neural processing routines can be used in conjunction with an extensive library of routines for image/signal processing and data visualization.

On a personal level, the work I did was interesting and relevant to my Ph.D. research, which was to identify features needed by a development environment for a visual programming language to aid in the design and implementation of modules for new application domains.

REFERENCES

- [1] S. A. Anderson and C. S. Williams, "Implementation of Neural Algorithms with the Khoros System," *Proceedings of the Workshop on Neural Networks: Academic/Industrial/NASA/Defense*, February 1992.
- [2] S. A. Anderson, W. J. Zimmermann, J. Schruben, A. S. Heger, T. L. Payne, "Neural Network Environment for Signal Processing and Pattern Recognition of Large Scale Data," *Proceedings of the Second Workshop on Neural Networks: Academic/Industrial/NASA/Defense*, February 1991, pp. 483-489.
- [3] B. Baird and F. Eeckman, "A Normal Form Projection Algorithm for Associative Memory," Internal document, University of California, Berkeley, CA, July 1992.
- [4] J. R. Rasure and C. S. Williams, "An Integrated Data Flow Visual Language and Software Development Environment," *Journal of Visual Languages and Computing*, Vol. 2, No. 3 (Sept. 1991), pp. 217-246.
- [5] J. A. Greenfield, "A Study of Oscillatory Neural Network Models," *Final Report: AFOSR Summer Research Program*, August 1992, pp. 10-1 - 10-20.

THE TRAJECTORIES OF EASTERN HEMISPHERE SOLAR ELECTRON
EVENTS
AS INFERRED FROM ISEE-3 RADIO AND PARTICLE DATA

Robert F. Willson
Research Associate Professor of Astronomy
Department of Physics and Astronomy

Tufts University
Medford, MA 02155

Final Report for:
Summer Research Program
Phillips Laboratory, Hanscom AFB

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

September 1992

THE TRAJECTORIES OF EASTERN HEMISPHERE SOLAR ELECTRON
EVENTS
AS INFERRED FROM ISEE-3 RADIO AND PARTICLE DATA

Robert F. Willson
Research Associate Professor
Department of Physics and Astronomy
Tufts University

Abstract

We use radio and particle data from the International Sun-Earth Explorer (ISEE-3) satellite to trace the paths of energetic electrons as they propagate from flare sites on the Sun to the earth. Our study extends over the period of operation of the electron experiment on ISEE-3 from August 1978 to November 1979. Previous investigations have suggested that particles in eastern hemisphere flares are much less likely to reach the terrestrial environment because they travel along Archimedean magnetic field spirals that curve away from the earth. Based on the timing of $H\alpha$ flares, and metric and kilometric Type III bursts, we first compiled a list of all electron events that could be associated with eastern hemisphere flares. The trajectories of the kilometric Type III bursts were then examined using data from the ISEE-3 radio experiment. Our results show that some of these associations were in fact spurious; the Type III bursts appeared to have originated from behind the west limb or from regions that are magnetically well-connected to the earth. Other, positively identified, eastern hemisphere bursts show complicated trajectories that suggest a role for coronal propagation or interactions with pre-existing structures that disrupt a well-ordered interplanetary field pattern.

I. Introduction

Observations of the Sun at X-ray wavelengths show that its outer atmosphere, or corona, consists of an inhomogeneous distribution of magnetic loops anchored to underlying sunspots. These structures evolve and occasionally interact resulting in the sudden release of magnetic energy both as electromagnetic radiation and as accelerated, energetic particles (electrons, protons and ions). Some of these particles are trapped in closed magnetic loops above the flare site, but others may escape and propagate into the earth's environment where they can produce geomagnetic storms, disrupt communications and interfere with high-flying aircraft or satellites.

Spacecraft observations have shown that most electron events are associated with flares in the Sun's western hemisphere (Lin et al. 1974) and this suggests that propagation effects control their characteristics. Electron streams are, in fact, believed to propagate along relatively narrow ($\theta \leq 30^\circ$) cones that follow pre-existing coronal and interplanetary magnetic fields (eg. Anderson et al. 1981) so that particles are observed only when the earth is within this cone of emission. The effects of solar rotation and the solar wind conspire to bend the interplanetary magnetic field into a well-defined Archimedean pattern which curves away from the earth-Sun direction for field lines that are anchored in the eastern solar hemisphere. Field lines that originate in the western hemisphere generally curve toward the earth, giving rise to so-called "well connected" regions on the solar surface. Energetic particles produced in western active regions are therefore more likely to be detected at the earth, while those from eastern hemisphere regions may miss the earth entirely.

Occasionally, electrons from eastern hemisphere flares are detected and these are thought to propagate great distances through the corona in order to be eventually released near the footpoints of field lines connecting to the earth. Because electrons do not easily diffuse across field lines in the interplanetary medium, it has been suggested that electron streams may propagate across low-lying coronal loops connecting different active regions on the solar surface (Wibberenz et al. 1989) or are later deflected by coronal mass ejections which disrupt the orderly spiral field pattern (Reiner and Stone 1986).

As the electrons propagate away from the flare site, they may also interact with the ambient plasma and excite Langmuir waves at the local plasma frequency ν_p . These plasma oscillations may in turn be converted to radio emission at the fundamental ($\nu = \nu_p$) or harmonic ($\nu = 2\nu_p$) of the plasma frequency giving rise to so-called Type III bursts. An association between Type III radio emission and energetic electrons ($E = \text{few} - 10\text{'s keV}$) has been firmly established (Lin 1970; Alvarez, Haddock and Lin 1972; Frank and Gurnett 1972; Lin Evans and Fainberg 1973) while intense electron plasma oscillations have also been observed in association with Type III bursts (Gurnett and Anderson 1976; 1977).

Type III bursts are characterized by a rapid drift from high to low frequency as

the electron streams encounter progressively lower densities and plasma frequencies farther from the Sun. Models of the electron density distribution in the interplanetary medium may therefore be combined with measurements of Type III burst frequency to infer the distance of electron streams from the Sun as a function of time. If electron streams from eastern hemisphere flares diffuse across coronal magnetic loops located close to the Sun, then the radio emission may appear to emanate from the western hemisphere. Alternatively, if the streams are deflected by coronal mass ejections or other pre-existing structures, then the Type III emission should exhibit irregular drifts in frequency as the electron streams encounter "kinks" or turns in the interplanetary magnetic field and plasma.

These effects may be investigated using radio and particle data from the International Sun-Earth Explorer (ISEE-3) which observed solar flares around the peak of the last solar activity cycle (1979-1982). In this paper we use data from the University of California, Berkeley energetic electron experiment (Anderson et al. 1978) to compile a list of all electron events which can be associated with eastern hemisphere flares. The flare associations are based on timing of $H\alpha$ flares and metric and decimetric Type III bursts and positional information as reported in Solar Geophysical Data. The paths of the electron streams associated with eastern hemisphere flares have been examined using data from the ISEE-3 radio experiment (Knoll et al. 1978). This approach allows us to visualize the large-scale magnetic field configurations through which the energetic electrons move and to therefore examine whether coronal propagation or interplanetary medium deflection scenarios apply.

II. Electron Events Associated With Eastern Hemisphere Flares

The University of California, Berkeley, experiment (Anderson et al. 1978) on board the ISEE-3 satellite was designed to measure the fluxes of interplanetary electrons with energies between 15 keV and ≈ 1 MeV. The data were recorded in 16 energy channels every 64 seconds and thereby providing measurements of the electron energy distribution as a function of the time.

A total of 326 impulsive electron events were detected by ISEE-3 during the 15 months of operation between August 1978 and November 1979 (Lin, private communication). In this study we have examined a subset of these events which, from a preliminary examination (B. Jackson, private communication), appear to be associated with eastern hemisphere flares. We have used $H\alpha$ flare and metric and decimetric Type III burst data tabulated in the *Solar Geophysical Data* monthly bulletins together with kilometric Type III burst data recorded by ISEE-3 to establish these identifications. Our criterion for a firm association was to select only those events for which a) a Type III burst occurred within ± 5 minutes of the $H\alpha$ flare onset and b) the $H\alpha$ flare and Type III burst began less than ≈ 30 minutes before the arrival of energetic electrons. A 30 min window corresponds to the approximate Sun-earth travel time for electrons

with an energies of a few tens of keV. Because these observations took place around the peak of solar activity, H α flares from widely-separated active regions (i.e., from eastern and western hemispheres) were often reported within the 30 minute interval prior to the electron event. These ambiguous events were therefore not included in our list of flare associations. We also rejected those events for which kilometric Type III events had been in progress prior to the H α flare or more than 5 minutes after the metric and decimetric type III burst. Decimetric Type III bursts, for example, are often associated with H α flares and are signatures of energetic electrons escaping along open magnetic fields above the flare site. The metric and decimetric and longer wavelength radio emission is thought to occur at the first or second harmonic of the local plasma frequency and because the coronal electron density decreases with height above the solar surface, an outwardly travelling disturbance will be observed first at the shorter radio wavelengths. The onsets of the metric and 1980 kHz ISEE-3 bursts (the highest frequency recorded by the spacecraft) should therefore occur at about the same time, and a time difference of more than a few minutes might therefore signal a coincidental association with an H α flare.

In Table 1 we list all electron events which fit our selection criteria. Here columns 1,2 and 3 give the date and onset time of the electron event, column 4 gives the start time of the 1980 kHz ISEE-3 burst, columns 5 and 6 give the start and end time of the metric Type III burst, and columns 7-11 give the start, maximum, location, importance and McMath active region number of the H α flare. We have also included a few flares in the western hemisphere (close to central meridian) because their particles are also normally expected to go undetected. As we will see in the next section a number of the Type III kilometric bursts have trajectories that are inconsistent with the location of the H α flare, and in fact suggest that the particles were produced by flares which occurred behind the west limb.

Table 1
Flare Associations of ISEE-3 Electron Events

Date	Electron Event		1980 kHz Onset	Metric		Optical Flare				Region
	Onset	Max		Onset	End	Onset	Max	Location	Imp.	
09/26/78	1400	1420	1340	1338	1348	1344	1345	N31 W04	-F	15551
10/05/78	1435	-	1405	1403	1413	1403	-	S18 E04	-N	15557
10/21/78	1630	1700	1558	1556	1602	1551	1620	N30 W12	1N	15598
11/08/78	1805	1830	1750	1751	1752	1751	1754	N18 E12	1B	15635
12/11/78	2105	-	2045	2044	2056	2045	2054	N23 E78	-N	15697
12/26/78	0650	0900	0810	0810	0811	0810	0814	N13 E34	1F	15733
01/11/79	2245	2380	2232	2230	2232	2227	2235	N17 E02	-N	15754
02/09/79	2210	2220	2145	2144	2126	2138	2210	N17 W09	-F	15807
03/11/79	0550	0650	0525	0525	-	0526	0531	S31 E26	-B	15869
04/14/79	0100	0120	0005	0001	0004	0006	0009	N21 E38	-F	15943
04/16/79	0620	0900	0525	0524	0533	0531	0534	N17 E16	1F	15952
04/28/79	1115	1145	1050	-	-	1049	-	S27 E21	-	15968
05/08/79	1600	2000	1515	1516	1517	1517	1517	N03 E50	-N	15995
05/20/79	1805	1830	1746	1744	1746	1746	1747	S20 E39	-F	16025
07/08/79	0200	0215	0050	0104	0152	0105	0105	N20 E22	-N	16124
07/17/79	1930	2000	1915	-	-	1913	1917	S20 E43	-N	16148
08/14/79	1738	1745	1700	1700	1732	1700	-	S27 E71	-F	16224
08/14/79	2100	2115	2055	2049	2052	2053	-	S22 E68	-B	16224
11/03/79	1500	1600	1440	1439	1445	1439	1443	S14 E17	-B	16401
11/04/79	2105	-	2045	2044	2056	2045	2054	N23 E78	-N	16414

III. Type III Burst Trajectories

The radio astronomy experiment on ISEE-3 was designed to provide high dynamic range observations of Type III bursts at 24 different frequencies between 24 kHz and 1980 kHz (see Knoll et al. 1978 for a description of the instrument and its capabilities). The spacecraft spins once every 3 seconds about an axis that is perpendicular to the ecliptic plane. Dipole antennas located in the ecliptic plane and along the spin axis allow the azimuth, elevation and approximate source size to be determined from an analysis of the signal's spin modulation (Fainberg 1979; Fainberg et al. 1985).

The source azimuth may be combined with an inferred radial distance from the Sun to calculate the trajectory of the Type III burst. As the flare electrons propagate outward from the Sun, they may radiate at the fundamental or harmonic of the plasma frequency, ν_p . By assuming a model for the interplanetary electron density it is therefore possible to calculate the heliocentric distance which corresponds to the emission at frequency ν or ν_p . The spatial location of the radio source can then be determined as the intersection point between this plasma sphere and the measured azimuth from ISEE-3.

The interplanetary electron density distribution, n_e , was assumed to be spherically symmetric and to fall off with heliospheric distance, R , as:

$$n_e(R) = 6.14R^{-2.10}$$

This relationship was derived by Bougeret et al. (1984) on the basis of *in-situ* observations of the solar wind by the Helios spacecraft. In terms of plasma frequency, $\nu_p = 9.1n_e^{1/2}$.

$$\nu_p(R) = 22.2R^{-1.05}$$

If we assume that the Type III emission occurs is the first harmonic (i.e., $\nu = 2\nu_p$), then the radial distance corresponding to a given ISEE-3 frequency is:

$$R(\nu) = 37.1 \nu^{-0.952}$$

where R and ν are expressed in units of AU and kHz, respectively. The frequency range of the ISEE-3 radiometers therefore corresponds to radio emission at heights of ≈ 0.27 to 1.5 AU from the Sun.

In Figure 1 we show an example of a kilometric burst observed by ISEE-3 on 26 September 1978. The time profile is characterized by an impulsive rise followed by a slower exponential decay. The peak of the burst occurs later at lower frequencies, as expected for an electron beam which passes through regions of progressively lower

density and plasma frequency in the interplanetary medium.

The trajectories for each of the Type III bursts listed in Table 1 were calculated from the ISEE-3 data using a fitting procedure that determines the source azimuth in each frequency channel. In general, the measurements are more reliable at the higher frequencies where the relatively small source sizes ($\theta, \approx 5^\circ$ at 1980 kHz) allow the spin modulation of the dipole antennas to locate the burst with greater accuracy. If the bursts were located at the radial distances implied by their plasma frequencies, then it should be straightforward to determine the source locations as the bursts travel between the earth and the Sun. For some frequencies, however, the radio bursts appear to lie at consistently greater radial distances than expected under the assumption that the emission is related to the plasma frequency. This implies that the electron density model assumed in equation 1 underestimates the true electron density along the measured azimuth. This is not surprising because Bougeret et al. (1984) have shown that the interplanetary electron density can vary by factors of up to 30 from that assumed in equation 1. For these frequencies, we have assumed that the source centroid lies tangent to the circle whose radius is perpendicular to the line of sight. In this case the radial distance is $R = \sin(AZ)$ AU.

In Figures 2 - 5 we show the trajectories of each of the Type III bursts (circles) listed in Table 1 together with the Archimedean spiral patterns of magnetic fields (dark lines) which project radially upward from the site of the $H\alpha$ flare. These predicted trajectories were calculated using the formulae given by Parker (1958) together with the solar wind speed data given by Cousens and King (1986). The filled circles represent source positions calculated from the density model of Bougeret et al. (1984), (equation 1), while the unfilled circles give the source position under the assumption that $R = \sin(AZ)$ AU. In general, the uncertainties in source azimuth become large ($\geq 10^\circ$ for $R \geq 0.5-0.75$ AU ($\nu \leq 100$ MHz)), and so these positions are not plotted.

IV. Discussion and Work for the Future

If the energetic electrons detected by ISEE-3 were injected radially up into the corona along field lines that connect to the site of the $H\alpha$ flare, and underwent no lateral diffusion through the corona (eg. across large-scale interconnecting coronal loops), the associated events we would expect that the associated Type III bursts would follow the simple Archimedean spiral trajectories through the interplanetary medium. Our plots of the measured Type III kilometric burst positions, however, show that this is not the case for many of these events. Some of these bursts, in fact, appear to originate from behind the west limb, or at least from longitudes that may be well-connected to the earth. Seven out of eighteen Type III bursts, in fact, appear to originate from behind the west limb, or at least from longitudes that are reasonably well connected to the earth.

The fact that nearly 40% of these flares/electron events seem to have been misiden-

tified makes one cautious about the conclusions of recent statistical studies which attempt to link electron events with specific flare longitudes without examining the Type III burst trajectories. Jackson and Leblanc (1989), for example, found that there was a peak in the number of flares between 50° and 70° west longitude associated with *in situ* electrons. This conclusion was based on examining associations between metric and kilometric Type III bursts and electron events *without* considering the trajectories of the ISEE-3 bursts. The implication is that at least some of Type III electron beams, to the extent that they can be detected by ISEE-3, are fairly limited ($\leq 30^\circ$) in angular extent.

Despite the fact that some of the events considered by Jackson and Leblanc may have been misidentified, we concur with the results of their study, namely that the detection of electrons at the earth is very unlikely when the flare is located in the eastern hemisphere or near central meridian. On the other hand, we have found a few interesting cases in which electrons *are* detected in association with Type III bursts that more or less follow Archimedean trajectories, and these must therefore require cones of emission that are considerably wider than 30° . Examples can be seen in the electron events of 1978 September 26, October 21, December 26, 1979 March 11, April 16, April 28, July 8, July 17 and November 4.

We have also found a few examples of trajectories that appear to bend towards the earth between $R = 0.1 - 0.5$ AU. The burst on 1978 October 5, for example, appears to deviate first to the east (at $R \approx 0.1$ AU), then back to the west (at $R \approx 0.28$ AU), thereafter following a trajectory that appears to intersect the earth. The event on 1978 November 8 also exhibits an interesting trajectory, namely an abrupt shift to the west near $R \approx 0.2$ AU, possibly indicating some interaction with a pre-existing structure at that distance.

We are currently examining some of the interesting ISEE-3 Type III trajectories in greater detail, in particular to calculate source positions with greater accuracy, both in azimuth and elevation. We will also examine other types of synoptic solar data, (such as records of soft X-ray burst from the GOES and SOLWIND satellites) which may give clues to why the paths of some events do not follow the Archimedean spiral pattern. Long duration soft X-ray bursts, for example, are often associated with coronal mass ejections which might disrupt well-ordered interplanetary magnetic fields through which the electron beams travel. The 1978 November 8 burst is also known to have taken place around the time of an event that produced an interplanetary shock, and this might explain why its trajectory appears to have changed abruptly, rather than gradually as for most of the other events.

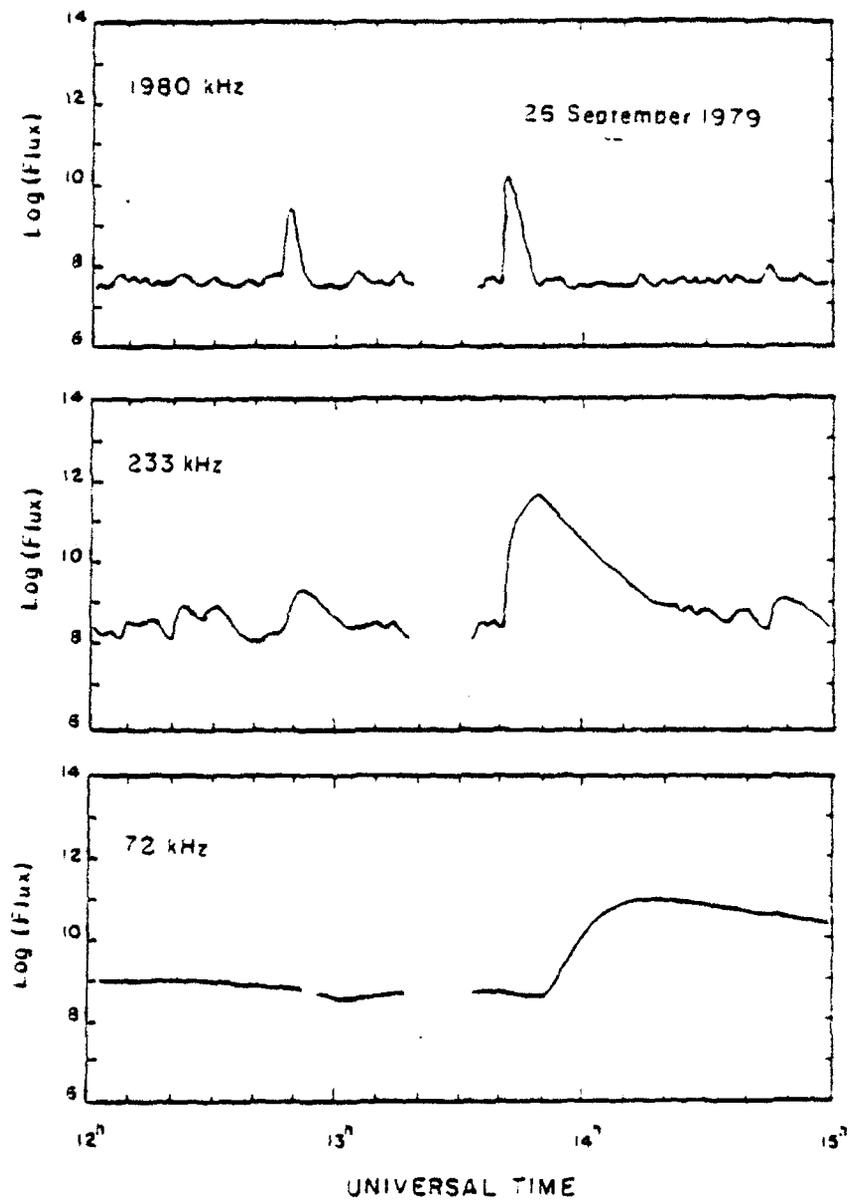
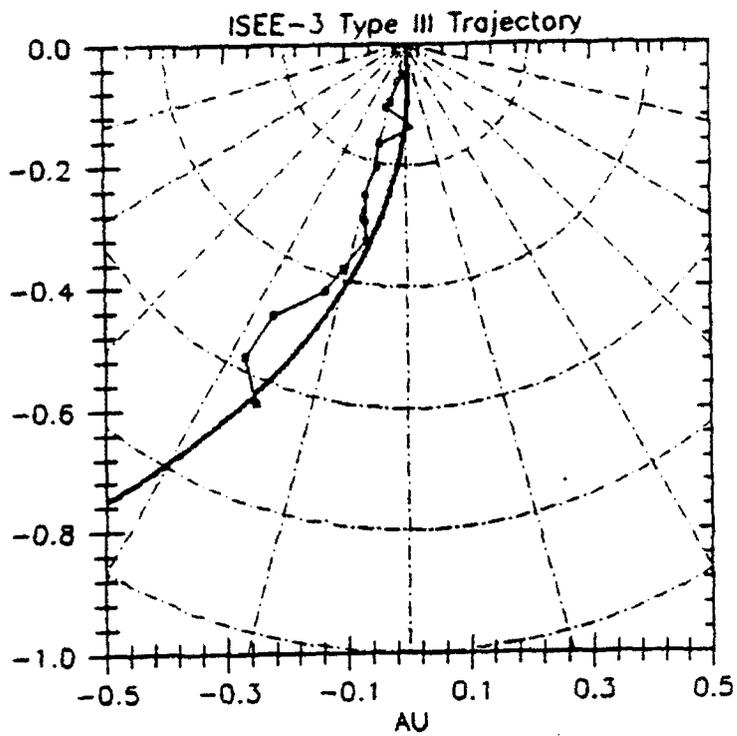
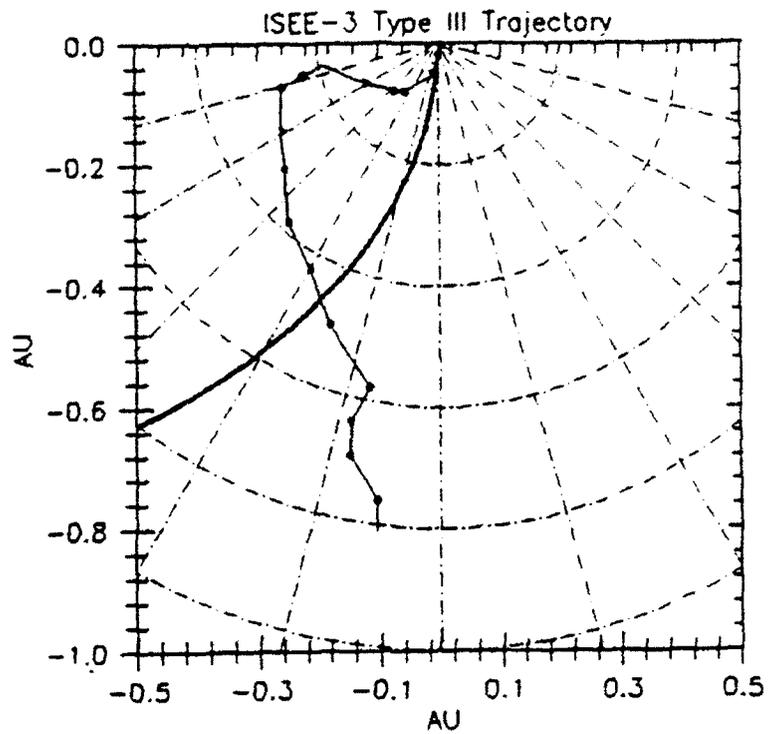


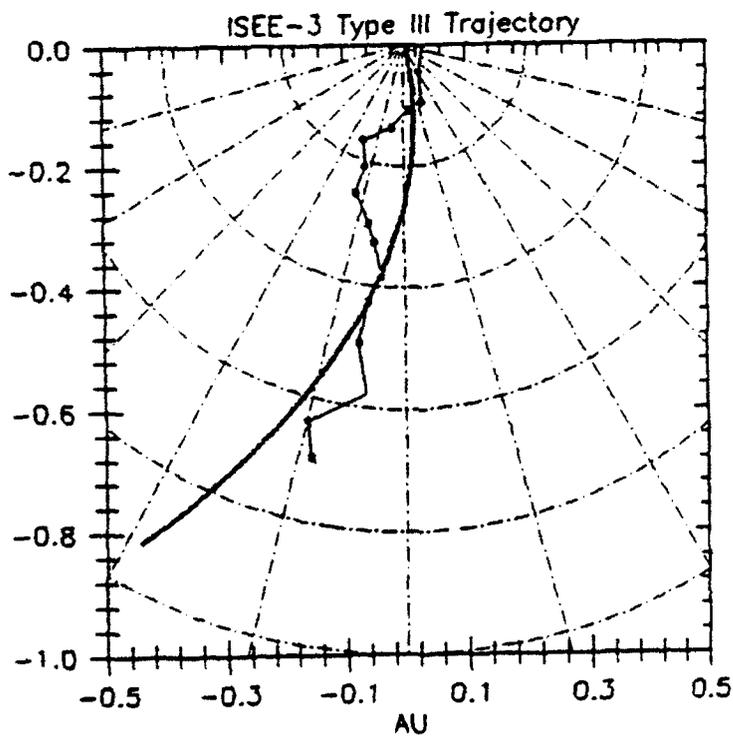
Figure 1



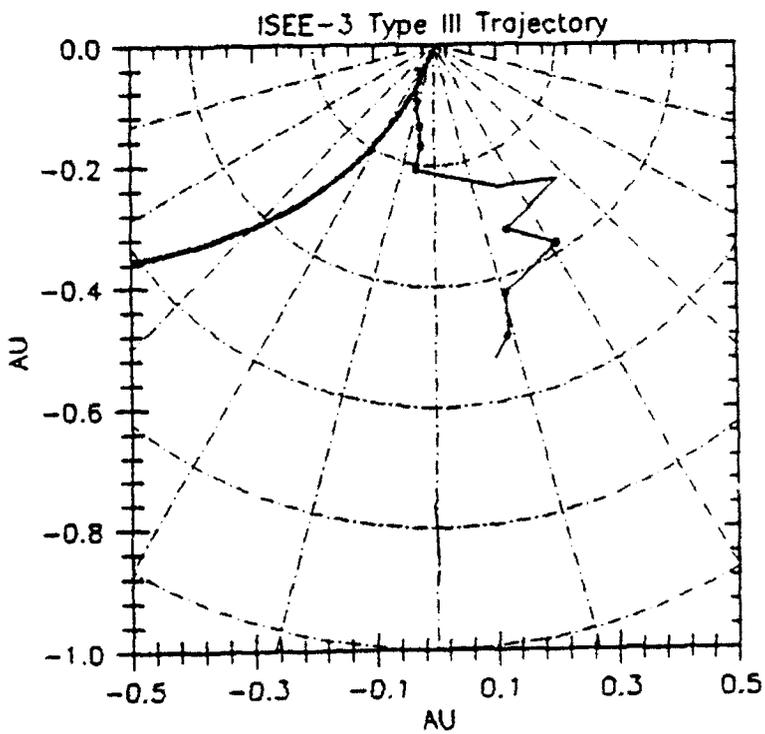
1978 September 26



1978 October 05

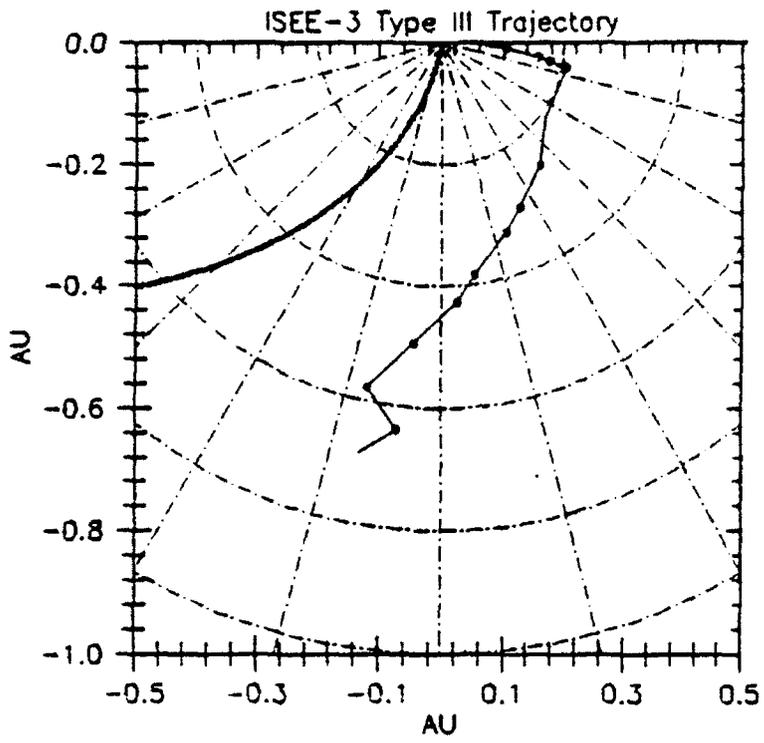


1978 October 21

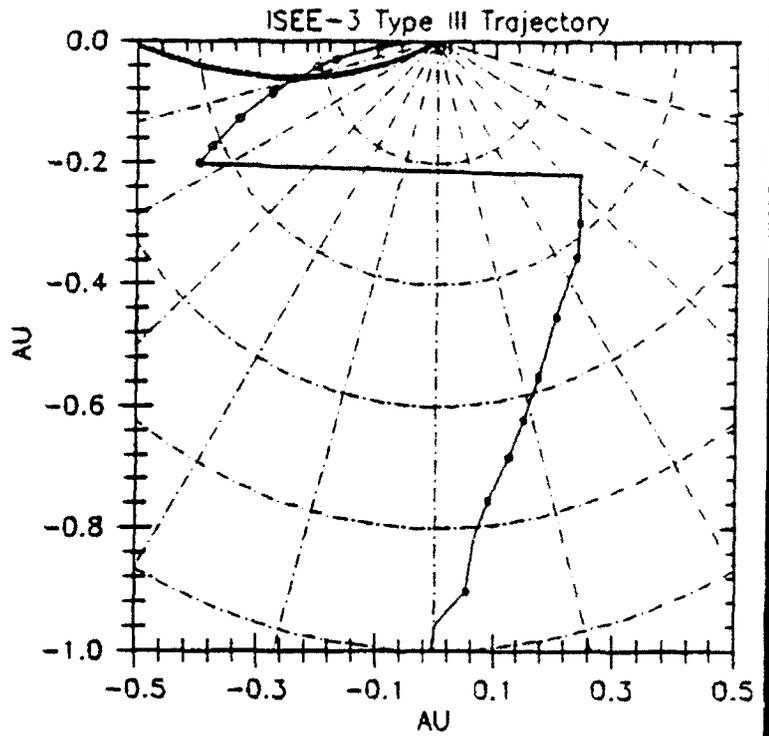


1978 November 08

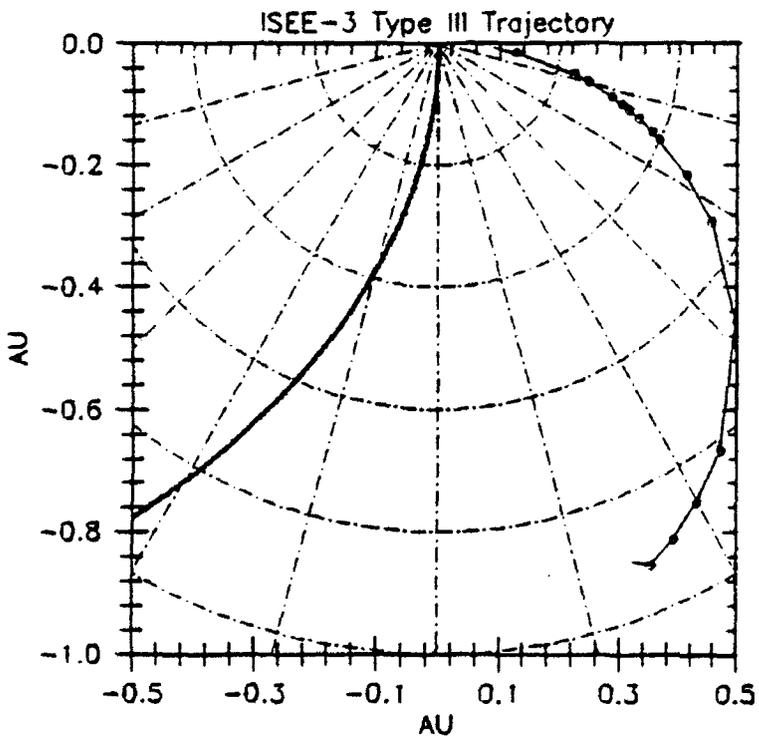
Figure 2



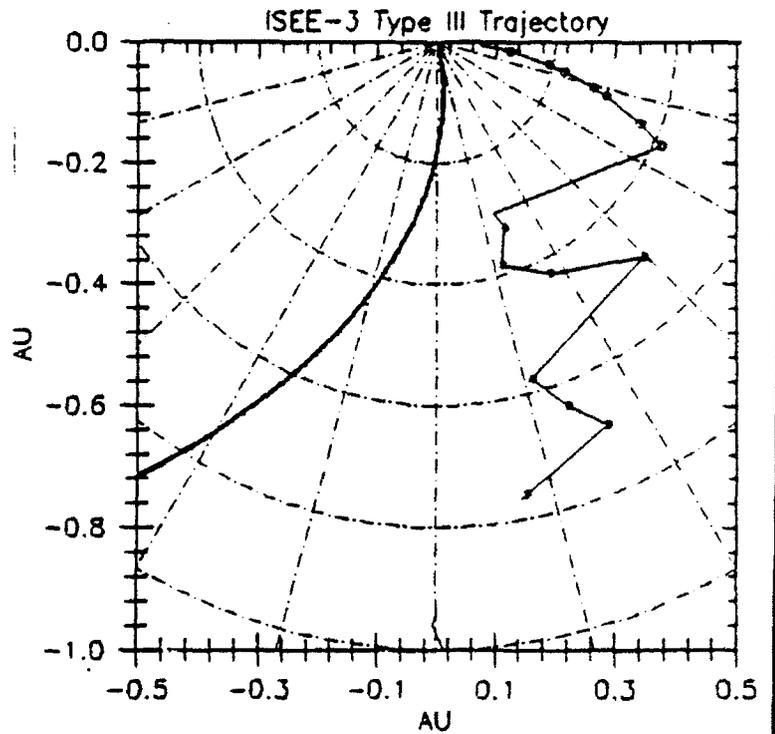
1978 December 11



1978 December 26



1979 January 11



1979 February 09

Figure 3

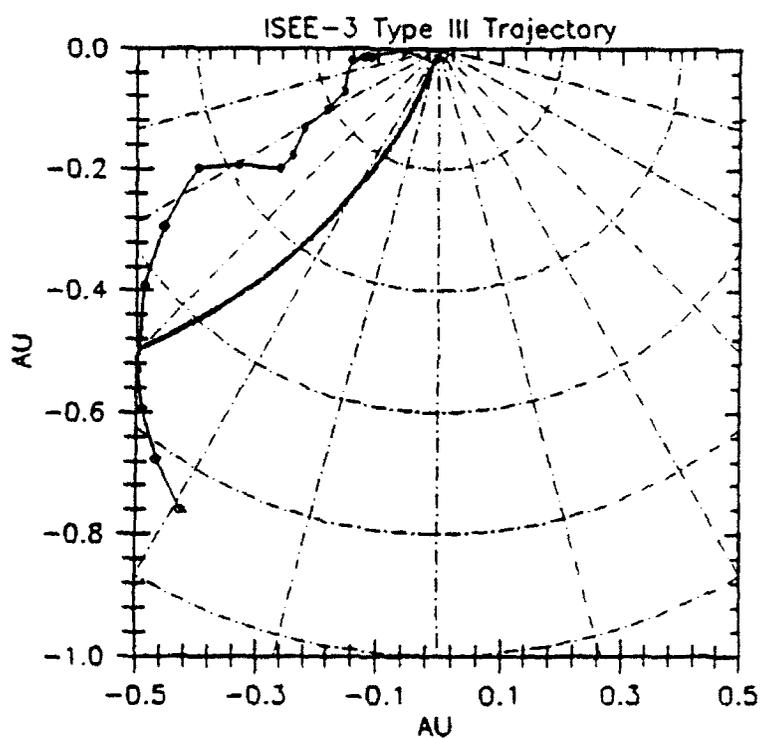
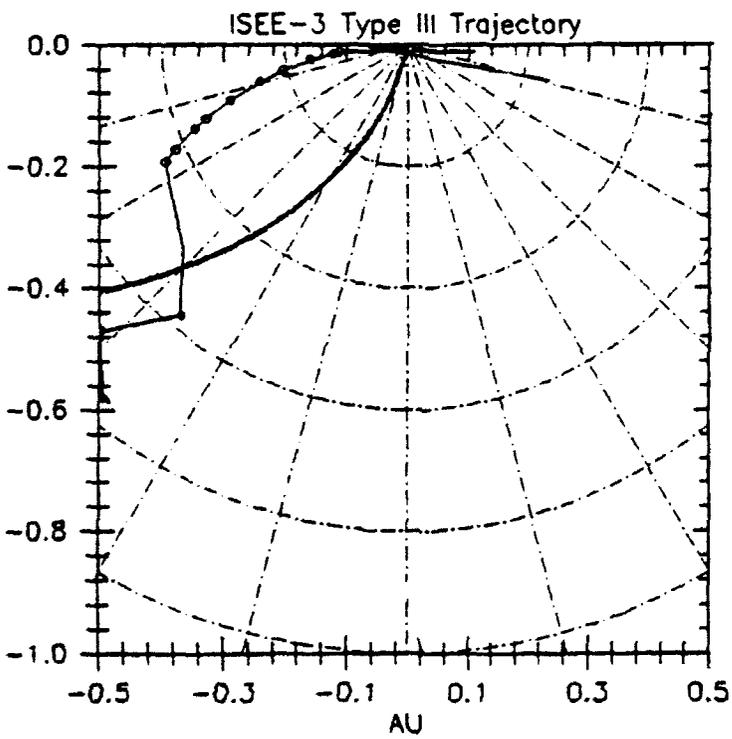
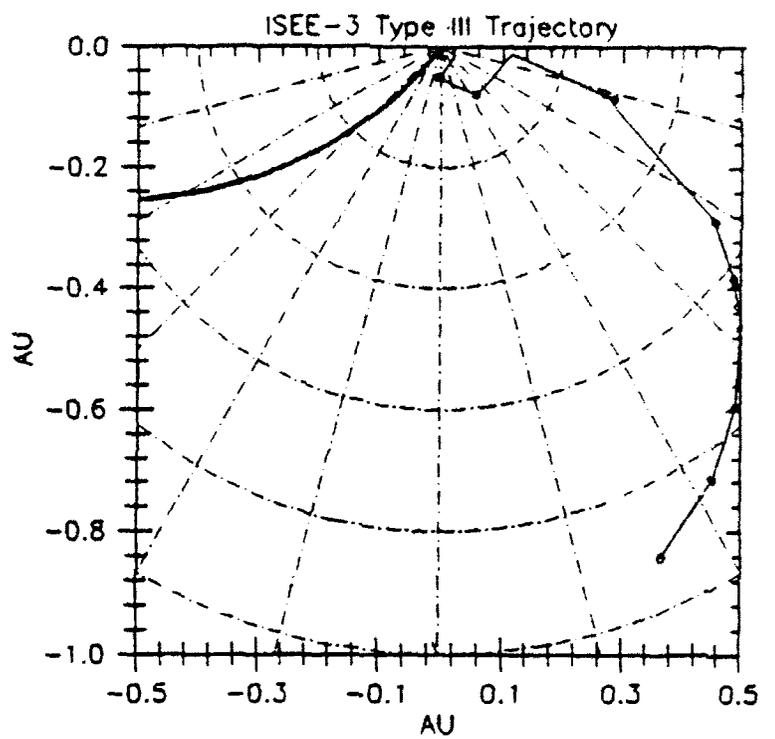
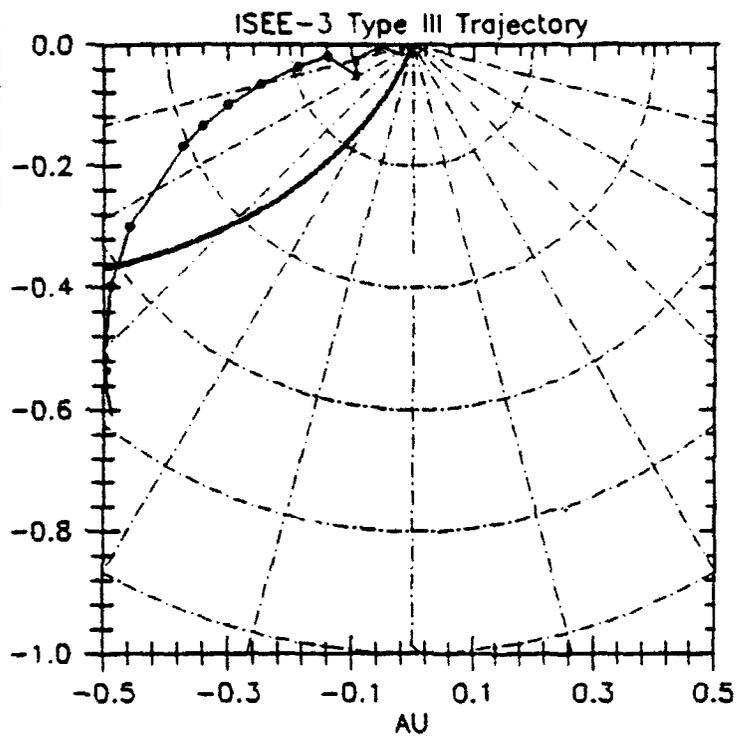
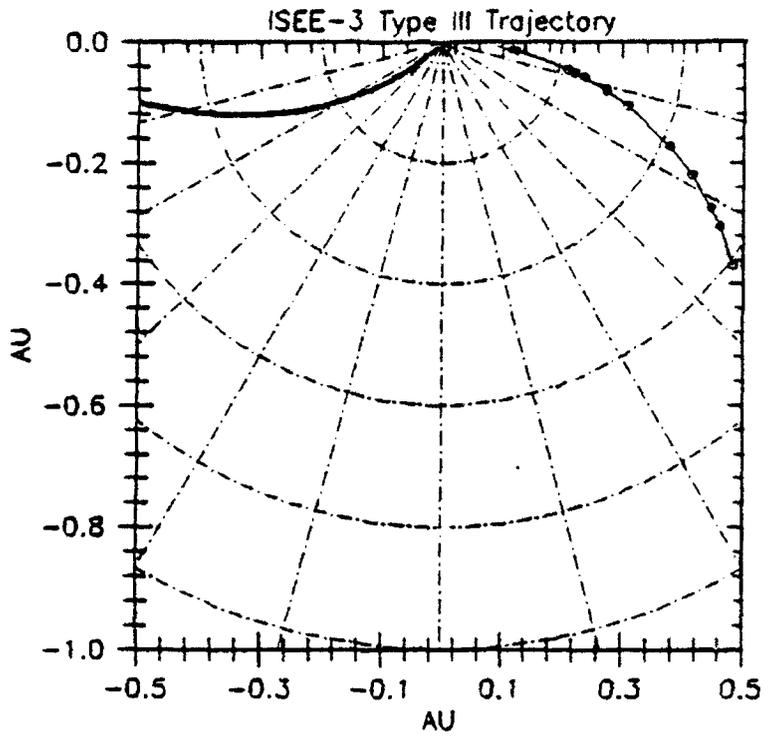
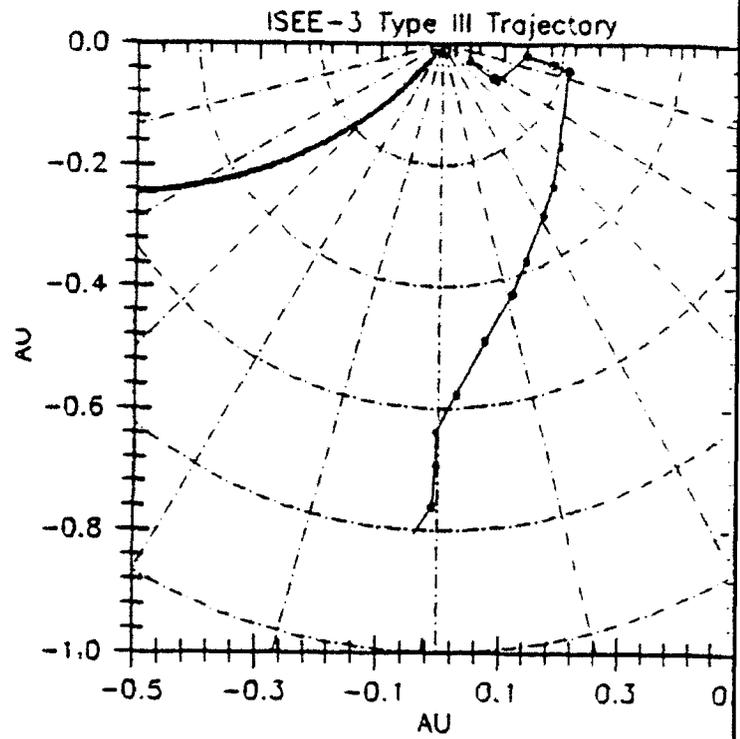


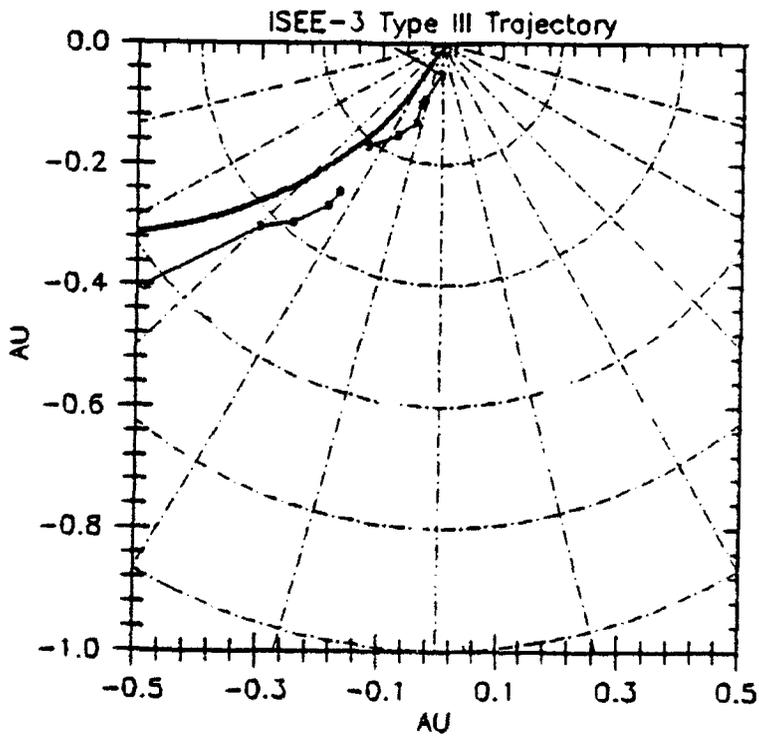
Figure 4



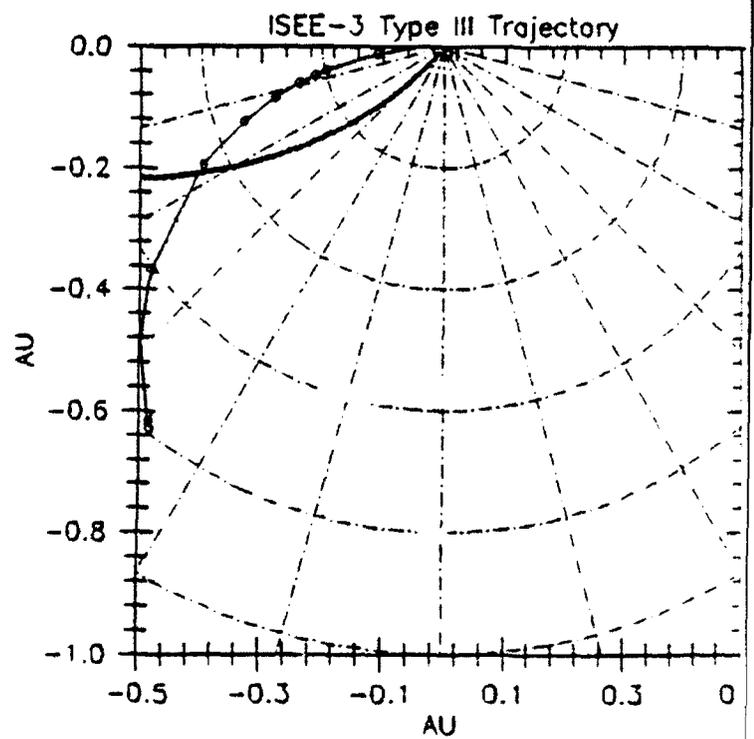
1979 May 08



1979 May 20



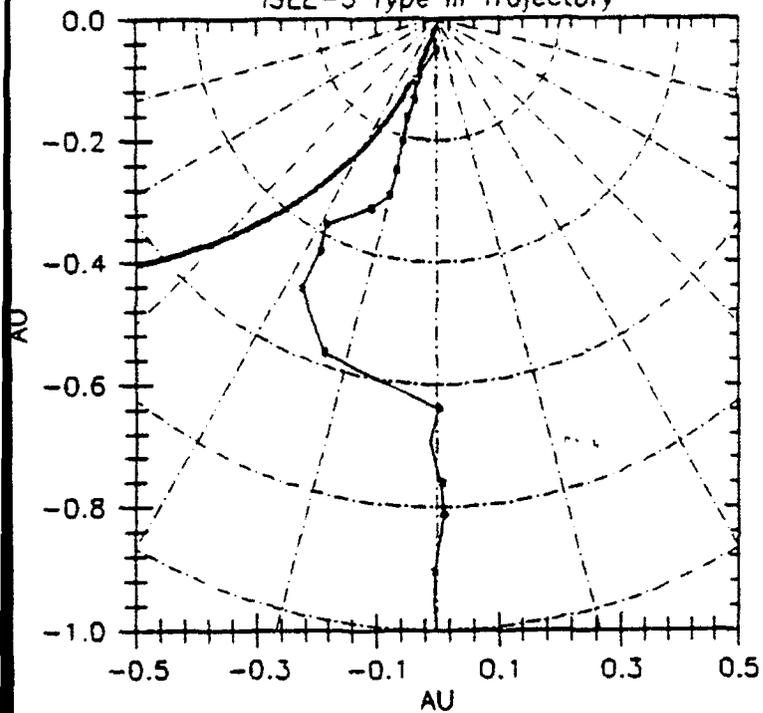
1979 July 08



1979 July 17

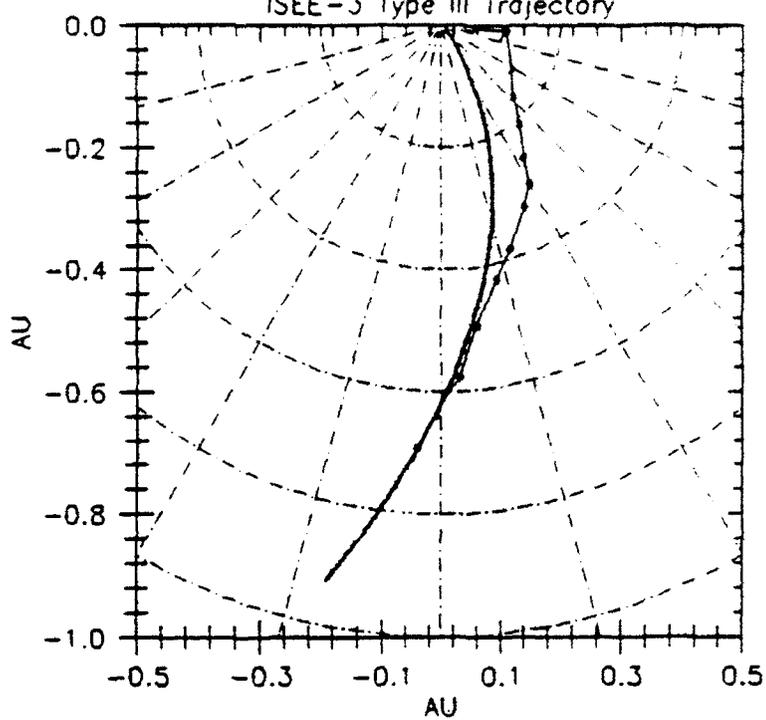
Figure 5

ISEE-3 Type III Trajectory



1979 November 03

ISEE-3 Type III Trajectory



1979 November 04

Figure 6

References

- Alvarez, H., Haddock, F.T., and Lin, R.P. 1972, *Solar Phys.*, **26**, 468.
- Anderson, K.K., McFadden, J.P., and Lin, R.P. 1981, *Geophys. Res letters*, **8**, 831.
- Anderson, K.A., Lin, R.P., Potter, D.W., and Heeterks, H.D. 1978, *IEEE Trans. Geosci. Elec*, **GE-16**, 153.
- Cousens, D.A., and King, J.H. 1986, in *Interplanetary Medium Data Book - Supplement 3*, NSSDC/WDC-R&S **86-04**.
- Fainberg, J. 1979, *NASA TM*, 80598.
- Fainberg, J., Huang, S., and Manning, R. 1985, *Astr. Ap.*, **153**, 145.
- Frank, L.A., and Gurnett, D.A. 1972, *Solar. Phys.*, **27**, 446.
- Gurnett, D.A., and Anderson, R.R. 1976, *Science*, **194**, 1159.
- Guernett, D.A., and Anderson, R.R. 1976, *J. Geophys. Res.*, **82**, 632.
- Jackson, B.V., and LeBlanc, Y. 1989, in M.A. Dubois, D. Gresollon and F. Bely Dubou, eds., *Plasma Phenomena in Solar Physics*. p.209.
- Knoll, R., et. al. 1978, *IEEE. trans. Geosci Electron.*, **GE-16**, 199.
- Lin, R.P. 1970, *Solar Phys.*, **12**, 266.
- Lin, R.P. 1974, *Space Rev.*, **16**, 189.
- Parker, E.N. 1988, *Ap. J.*, **128**, 664.
- Reiner, M.J., and Stone, R.G. 1986, *EOS*, **67**, 326.
- Wibberenz, G. et al. 1989, *Solar Phys.*, **124**, 353.